NON-PARAMETRIC INFERENCE AND REGRESSION ANALYSIS FOR CUMULATIVE INCIDENCE FUNCTION UNDER TWO-STAGE RANDOMIZATION

by

Idil Yavuz

B. S., Dokuz Eylul University, 2005

M. A., University of Pittsburgh, 2012

Submitted to the Graduate Faculty of

the Department of Statistics in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

THE DIETRICH SCHOOL OF ARTS & SCIENCES

This dissertation was presented

by

Idil Yavuz

It was defended on

June 4th 2013

and approved by

Yu Cheng, Assistant Professor, Departments of Statistics and Psychiatry

Satish Iyengar, Professor, Department of Statistics

Allan R. Sampson, Professor, Department of Statistics

Abdus S. Wahed, Associate Professor, Department of Biostatistics

Dissertation Director: Yu Cheng, Assistant Professor, Departments of Statistics and Psychiatry

NON-PARAMETRIC INFERENCE AND REGRESSION ANALYSIS FOR CUMULATIVE INCIDENCE FUNCTION UNDER TWO-STAGE RANDOMIZATION

Idil Yavuz, PhD

University of Pittsburgh, 2013

In recent years, personalized medicine and dynamic treatment regimes have drawn considerable attention and two-stage randomization is commonly used to gather data for making inference on dynamic treatment regimes. Meanwhile, more and more practitioners become aware of competing risk censoring, where subjects are exposed to more than one possible failure and the event of interest may be dependently censored by the occurrence of competing events.

We aim to compare several treatment regimes from a two-stage randomized trial on survival outcomes that are subject to competing-risk censoring. With the presence of competing risks, cumulative incidence function (CIF) has been widely used to quantify the cumulative probability of occurrence of the target event by a specific time point.

In the first part of this dissertation, we propose non-parametric estimators for the CIF using inverse weighting, and provide inference procedures based on the asymptotic linear representation to help compare the CIFs from two different treatment regimes. Through simulation, we show the practicality and advantages of the proposed estimators and apply them to data from the Cancer and Leukemia Group B (CALGB) trial.

Next, we propose a pattern-mixture type estimator for the CIF. Pattern-mixture models stratify data according to dropout patterns, make estimates of a certain parameter on each stratum, and obtain the final estimate by taking a weighted average of these estimates.

We show that this approach can be borrowed for estimating the CIF under a two-stage randomization. We investigate its properties using simulation and apply it to the CALGB data.

In the third part, we focus on regression analysis under a two-stage randomization setting. Even though extensive research is being carried out by researchers on the regression problem for dynamic treatment regimes, no research has been done on modeling the CIF when a two-stage randomization has been carried out. We extend the multi-state (Cheng et al., 1998), Fine and Gray (1999) and Scheike et al. (2008) regression models for modeling the CIF of dynamic treatment regimes and provide ways to implement the proposed models in R using the existing packages. We show the improvement our methods provide by simulation.

Keywords: competing risks, cumulative incidence function, dynamic treatment regime, inverse weighting, multi-state model, pattern-mixture models, proportional hazards for subdistribution, regression analysis, time-varying effects, two-stage randomization.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Dynamic treatment regime is a set of rules that guides treatment according to the patients' needs, observed covariates and intermediate responses. The idea is to find the regime that in the end produces the best outcome, where the best outcome can be the longest survival time or the biggest reduction in some pre-specified score. To achieve this goal, two-stage randomization designs have been used by researchers and have become more popular in cancer, AIDS and psychiatric studies. Designing such studies and methods for analyzing data that arise from these designs have been greatly studied in recent years.

In this dissertation, we focus on analyzing the data from two-stage randomization trials for a survival outcome that is subject to competing-risk censoring. Competing-risk censoring refers to a situation where subjects in a study are exposed to more than one possible failure and the specific event of interest may be dependently censored by the occurrence of competing events. Competing-risk censoring is more natural than assuming the censoring due to other events is independent, as a result this type of setting is gaining more attention from practitioners when analyzing time-to-event data (Koller et al., 2012). When a subject is exposed to more than one risk, we are then interested in the probability of occurrence of the event of interest at a specific time point, or the cumulative incidence function (CIF). This quantity is intuitively interpretable and non-parametrically identifiable, hence it has been commonly used in the competing risks literature (Kalbfleisch and Prentice, 2002). When competing risks are present and a two-stage randomization design is being used, the objective then would become finding a regime which results in a reduced probability of occurrence of the event of interest. Although methods for analyzing time-to-event data that arise from two-stage randomization designs have been developed when there is only one cause for fail-

ure, no research has been carried out on how to analyze data from such designs under a competing risks setting.

In this dissertation, we will first propose inverse weighted estimators that can be used for estimating the CIF of dynamic treatment regimes. We will then give inference procedures that can be used for comparing two regimes. Through simulation we will show that the proposed estimators are unbiased and they all out-perform the standard estimator naïvely applied to data from subjects consistent with a specific regime.

In the second part of this dissertation, we propose another estimator for the CIF which can be used when dynamic treatment regimes are of concern, following ideas of pattern-mixture models. This pattern-mixture estimator is straightforward with a simple explicit variance estimator. We show the unbiasedness and good performance of this estimator through simulation.

Finally, we will examine the situation where covariates are involved and when competing risks are present. It is of practical interest to examine how covariates may affect the CIF for a specific regime. Although several regression models have been proposed for estimating the CIF, none can be directly applied in a two-stage randomization design. We will propose two methods to improve current models and provide extensive simulation results.

2.0 NON-PARAMETRIC INFERENCE FOR THE CUMULATIVE INCIDENCE FUNCTION UNDER TWO-STAGE RANDOMIZATION

2.1 INTRODUCTION

In dynamic treatment regimes, the treatment level and type can vary depending on evolving measurements of subject-specific need for treatment. These regimes are rules that provide treatment adapted to individual needs, thus they are more cost effective and they improve patient's compliance by avoiding over-treatment or under-treatment (Lavori and Dawson, 2000). Different designs of clinical trials can be used for estimating the treatment efficacy of various treatment regimes. The simplest one is the single stage randomization design where the patients are randomized to all possible treatment regimes upon entry into the trial. However, this method is not cost effective and in general requires a larger sample size. The second way is to conduct multiple separate trials for the first and second stages but this raises issues about patient comparability between trials. A better design is the so-called Sequential Multiple Assignment Randomized Trial (SMART) which was considered by Lavori and Dawson (2000), Lavori and Dawson (2004), Murphy (2005) and Murphy et al. (2007), where patients are randomized to the initial treatment options at entry and those continuing to the next stage are randomized to available treatment options based on their intermediate response to the initial treatment and randomization is continued in this fashion. For example, in a two-stage randomization design, suppose there are two treatment options A_1 and A_2 at the first stage, and two treatment options for both responders and non-responders namely B_1 and B_2 and $B_1^{'}$ and $B_2^{'}$ at the second stage. Using the SMART strategy the randomization can be carried out as shown in Figure1.

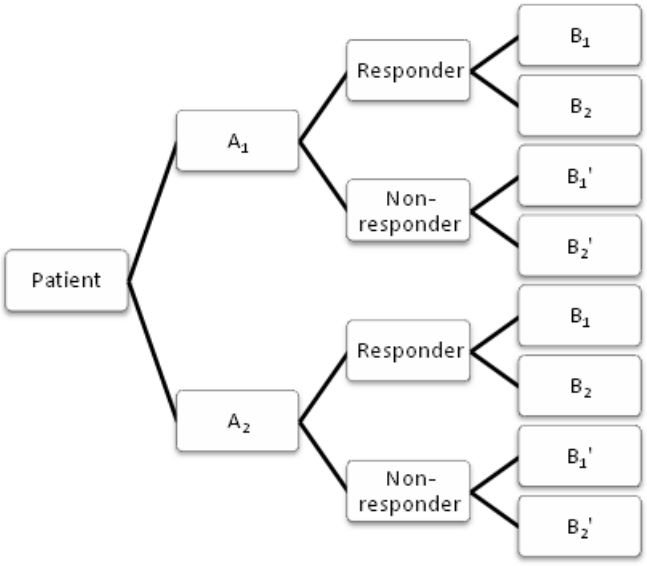

Figure 1: A two-stage randomization setup

Statistical methods for analyzing data from a SMART design are available in the literature. For example, Murphy et al. (2001) developed a marginal mean model for the mean response of a dynamic treatment regime after which they provided a methodology for constructing an optimal regime (Murphy, 2003). Lunceford et al. (2002) used the inverse weighting method introduced by Robins et al. (1994) to propose a marginal mean model (Murphy et al., 2001) for analyzing survival data from a two-stage setting. Later, Guo and Tsiatis (2005) proposed a weighted risk-set estimator which is a modified Nelson-Aalen estimator (Aalen, 1978). Although it may seem appropriate at first to apply the Kaplan-Meier estimator (Kaplan and Meier, 1958) to the subgroup of patients following a specific regime, Wahed and Tsiatis (2006) showed that such an estimator is biased, and to correct for this bias they proposed a weighted version of the Kaplan-Meier estimator following Lunceford et al. (2002) and Lokhnygina and Helterbrand (2007). Murphy and Bingham (2009) used screening experiments to help develop dynamic treatment regimes.

Although the inference for dynamic treatment regimes has been studied in various articles until now there has been no research on how to estimate the survival time for these regimes under competing risks settings. Competing-risk censoring is common in practice when there

are composite outcomes, and it has recently drawn more attention from practitioners (Gooley et al., 1999; Klein, 2006; Koller et al., 2012). In competing risks settings, the cumulative incidence function (CIF) is a commonly used quantity which describes the proportion of the event of interest occurring over time in the presence of competing events (Kalbfleisch and Prentice, 2002).

In this chapter, we show that the standard non-parametric estimator of the CIF, using the data from only those subjects who are consistent with the treatment regime of interest, may be biased. Instead, we propose several estimators of the CIF with various weight functions, and compare them to the standard non-parametric estimator. After introducing the necessary notation, defining the new estimators and proposing inference for comparison of two regimes, we will present the findings from sets of simulations we ran and finally demonstrate the practicality of the proposed methods by applying them to the CALG-B data set.

2.2 SET-UP AND NOTATION

The design in Figure 1 would create a total of eight regimes $A_jB_kB_l^{'}$ for j,k,l=1,2. Here $A_jB_kB_l^{'}$ stands for the regime where the subject is treated with A_j followed by B_k if the subject responds to A_j and by $B_l^{'}$ if not. Let T_{jkl} denote the survival time of a subject following the regime $A_jB_kB_l^{'}$. The goal is then to estimate the probability that the event of interest occurs before a specific time t where the subject may fail from any of the competing events of failure. The survival distribution of a subject following the regime $A_jB_kB_l^{'}$ is $S_{jkl}(t) = P(T_{jkl} > t)$ but estimating this overall survival will not provide detailed information about the probabilities of interest thus the estimator recommended for this purpose is the CIF.

Without loss of generality, we consider only the subjects that are assigned to treatment A_1 at first stage and assume there are only two causes of possible failures. It is easy to show that adding extra layer of weight to account for those subjects who are initially assigned to

A_2 will not change the inference procedure, if the weight does not depend on covariates. For the case where the weight does depend on covariates, Miyahara and Wahed (2010) pointed out that there may not be much gain in efficiency in the inference of a survival outcome with independent censoring, since covariates are most likely balanced due to randomization. In addition, multiple competing events can be grouped together without affecting the analysis for the event of interest. Therefore, limiting the number of causes of possible failure does not affect the inference proposed.

For the i^{th} subject ($i = 1, \cdots, n$), let T_i^R be the time to intermediate response assessment since the initial randomization, R_i be the response indicator (R_i=1 if the subject has responded to A_1; 0 otherwise), Z_{1i} be the second treatment assignment indicator for responders ($Z_{1i} = k$ if subject is assigned to B_k; $k = 1, 2$), and Z_{2i} be the second treatment assignment indicator for non-responders ($Z_{2i} = l$ if subject is assigned to $B_l^{'}$; $l = 1, 2$). Let T_i denote the time to first event since the initial randomization and ϵ_i denote the corresponding event type (=1, if the first event is the event of interest, =2, if the competing event occurs first).

There may also be independent censoring present which can be written as C_i. Hence the observed event time is $V_i = \min(T_i, C_i)$, and the cause indicator $\Delta_i = \epsilon_i I(C_i \geq T_i)$ takes on the value of 1 or 2 if the cause 1 or 2 event occurs before censoring, and 0 if no event is observed before C_i. Then, the i^{th} subject's data can be represented as $\{T_i^R, R_i, R_i Z_{1i}, (1 - R_i)Z_{2i}, V_i, \Delta_i\}$. Here, T_{jkl} is only observed for subjects who were on treatment A_j, responded to it and received B_k or did not respond to A_j and received $B_l^{'}$, and not observed for the others. The randomization probabilities $\pi_{B_k} = Pr(Z_{1i} = k \mid R_i = 1)$ and $\pi_{B_l^{'}} = Pr(Z_{2i} = l \mid R_i = 0)$ are assumed to be independent of the observed data prior to the second randomization except for R_i. In some cases, the time to response may also be censored but in such cases it is customary to treat the patients with censored response times as non-responders (Lunceford et al., 2002).

2.3 THE WEIGHTED ESTIMATORS OF THE CIF

2.3.1 The CIF Estimator with Fixed Weights

To estimate the cause 1 CIF for the regime $A_1B_kB_l^{'}$, $k,l=1,2$, one may naïvely construct a standard non-parametric estimator using the data only from those subjects whose treatments are consistent with the regime (i.e., subjects who were on treatment A_1, responded to it and received B_k or did not respond to A_1 and received $B_l^{'}$). Let $t_1<t_2<\cdots<t_k$ be the distinct event times where either the event of interest or the competing event occurs. Let Y_i be the number of subjects at risk, d_i be the number of subjects with the occurrence of the event of interest, and r_i be the number of subjects with the occurrence of the competing event at time t_i among patients who received B_k or $B_l^{'}$ as second stage therapy. Then, the cause 1 CIF for the regime $A_1B_kB_l^{'}$ would be estimated by

$$\hat{F}_{1,A_1B_kB_l^{'}}(t)=\sum_{t_i\leq t}\frac{d_i}{Y_i}\left\{\prod_{j=1}^{i-1}\left(1-\frac{d_j+r_j}{Y_j}\right)\right\} \tag{2.3.1}$$

for $t_1\leq t$ and 0 otherwise. For $t_1\leq t$ the CIF can be represented as

$$\hat{F}_{1,A_1B_kB_l^{'}}(t)=\sum_{t_i\leq t}\hat{S}_{A_1B_kB_l^{'}}(t_i-)\frac{d_i}{Y_i},$$

where $\hat{S}_{A_1B_kB_l^{'}}(t_i-)$ is the Kaplan-Meier estimator (Kaplan and Meier, 1958) evaluated at just before time t_i. The variance estimator of $\hat{F}_{1,A_1B_kB_l^{'}}(t)$ is given in Klein and Moeschberger (2003) as

$$\begin{aligned}\hat{\sigma}^2\left(\hat{F}_{1,A_1B_kB_l^{'}}(t)\right) &= \sum_{t_i\leq t}\Bigg(\hat{S}_{A_1B_kB_l^{'}}(t_i)^2\left\{\left(\hat{F}_{1,A_1B_kB_l^{'}}(t)-\hat{F}_{1,A_1B_kB_l^{'}}(t_i)\right)^2\frac{r_i+d_i}{Y_i^2}\right\}\\ &\quad+\hat{S}_{A_1B_kB_l^{'}}(t_i)^2\left[\left\{1-2\left(\hat{F}_{1,A_1B_kB_l^{'}}(t)-\hat{F}_{1,A_1B_kB_l^{'}}(t_i)\right)\right\}\frac{d_i}{Y_i^2}\right]\Bigg).\end{aligned}$$

The standard CIF estimator discards all the information from those subjects who are not consistent with the regime $A_1B_kB_l^{'}$, hence it loses efficiency and may be biased. To account for loss of those subjects, we take a similar approach used by Lunceford et al. (2002), Guo

and Tsiatis (2005) and Miyahara and Wahed (2010), and propose a weighted cumulative incidence function (WCIF) estimator

$$\hat{F}^{w}_{1,A_1B_kB_l'}(t) = \sum_{t_i \le t} \frac{d_i^w}{Y_i^w} \left\{ \prod_{j=1}^{i-1} \left(1 - \frac{d_j^w + r_j^w}{Y_j^w} \right) \right\}, \tag{2.3.2}$$

for $t_1 \le t$ and 0 otherwise, where $d_i^w = \sum_{j=1}^n I(V_j = t_i, \Delta_j = 1) Q_{A_1B_kB_l',j}$, $r_i^w = \sum_{j=1}^n I(V_j = t_i, \Delta_j = 2) Q_{A_1B_kB_l',j}$ and $Y_i^w = \sum_{j=1}^n I(V_j \ge t_i) Q_{A_1B_kB_l',j}$ for $Q_{A_1B_kB_l',j} = R_j I\{Z_{1j} = k\}/\pi_{B_k} + (1 - R_j) I\{Z_{2j} = l\}/\pi_{B_l'}$. Recall that Z_{1j} is the second treatment assignment indicator for responders ($Z_{1j} = k$ if subject j is assigned to B_k; $k = 1, 2$), Z_{2j} is the second treatment assignment indicator for non-responders ($Z_{2j} = l$ if subject j is assigned to B_l'; $l = 1, 2$), $\pi_{B_k} = P(Z_{1j} = k \mid R_j = 1)$ and $\pi_{B_l'} = P(Z_{2j} = l \mid R_j = 0)$. The CIF estimator in (2.3.2) is similar to the standard estimator except that those subjects following the regime are inversely weighted by the probability of being allocated to a specific treatment option during the second stage to compensate for those subjects who have been assigned to alternative treatments but could have been consistent with the regime if there had been no second randomization.

To estimate the variance of $\hat{F}^{w}_{1,A_1B_kB_l'}$, the following counting process formulation was used. For subject j, define the weighted cause specific event processes $N_{1j}^w(s) = I(V_j \le s, \Delta_j = 1) Q_{A_1B_kB_l',j}$ and $N_{2j}^w(s) = I(V_j \le s, \Delta_j = 2) Q_{A_1B_kB_l',j}$, and the overall event process $N_j^w(s) = N_{1j}^w(s) + N_{2j}^w(s)$. Also define the weighted at-risk process $Y_j^w(s) = I(V_j \ge s) Q_{A_1B_kB_l',j}$. Summing over all subjects, we have $Y_\cdot^w(s) = \sum_{j=1}^n Y_j^w(s)$, $N_{1\cdot}^w(s) = \sum_{j=1}^n N_{1j}^w(s)$ similarly $N_{2\cdot}^w(s)$, and $N_\cdot^w(s) = N_{1\cdot}^w(s) + N_{2\cdot}^w(s)$. Let $\Lambda_1(s) = \int_0^s \lambda_1(u) du$, where $\lambda_k(u) = \lim_{h \to 0} \frac{P(u \le V < u+h, \Delta = k \mid V \ge u)}{h}$ is the cause-specific hazard function for event k, and $\Lambda(s) = \int_0^s \lambda(u) du$, where $\lambda(u) = \lim_{h \to 0} \frac{P(u \le V < u+h \mid V \ge u)}{h}$ is the all-cause hazard. Let $M_j^w(s) = N_j^w(s) - \int_0^s Y_j^w(u) d\Lambda(u)$. One can show that M_j^w's are martingales and so is $M_\cdot^w(s) = N_\cdot^w(s) - \int_0^s Y_\cdot^w(u) d\Lambda(u)$. Similarly, $M_{1\cdot}^w(s) = N_{1\cdot}^w(s) - \int_0^s Y_\cdot^w(u) d\Lambda_1(u)$ is also a martingale. Using the counting process notation the weighted survival and the WCIF estimator can be represented as follows:

$$\hat{S}^{w}_{A_1B_kB_l'}(t) = \prod_{s \le t} \left\{ 1 - \frac{\Delta N_\cdot^w(s)}{Y_\cdot^w(s)} \right\},$$

$$\hat{F}^w_{1,A_1B_kB_l'}(t) = \int_0^t \hat{S}^w_{A_1B_kB_l'}(s-)d\hat{\Lambda}^w_1(s),$$

where $\Delta N^w_{\cdot}(s) = N^w_{\cdot}(s) - N^w_{\cdot}(s-)$, and $\hat{\Lambda}^w_1(s) = \int_0^s \frac{dN^w_{1\cdot}(s)}{Y^w_{\cdot}(s)}$.

To simplify the notation in what follows, we will temporarily drop the regime from the notation by letting $\hat{F}^w_{1,A_1B_kB_l'} = \hat{F}^w_1$ and let $\hat{S}^w_{A_1B_kB_l'} = \hat{S}^w$. To derive the variance of the WCIF estimator, we write $\sqrt{n}(\hat{F}^w_1(t) - F_1(t))$

$$\begin{aligned} &= \sqrt{n}\left\{\int_0^t \hat{S}^w(s-)d\hat{\Lambda}_1(s) - \int_0^t S(s-)d\Lambda_1(s)\right\} \\ &= \sqrt{n}\left\{\int_0^t \hat{S}^w(s-)\frac{dN^w_{1\cdot}(s)}{Y^w_{\cdot}(s)} - \int_0^t \hat{S}^w(s-)d\Lambda_1(s) + \int_0^t (\hat{S}^w(s-) - S(s-))d\Lambda_1(s)\right\}. \end{aligned} \quad (2.3.3)$$

Following Theorem 3.2.3 in Fleming and Harrington (1991) the ratio of the weighted survival function and the survival function can be derived as follows:

$$\begin{aligned} \frac{\hat{S}^w(t)}{S(t)} &= \frac{\hat{S}^w(0)}{S(0)} - \int_0^t \frac{\hat{S}^w(s-)}{S(s)S(s-)}dS(s) + \int_0^t \frac{1}{S(s)}d\hat{S}^w(s) \\ &= 1 + \int_0^t \frac{\hat{S}^w(s-)}{S(s)}d\Lambda(s) - \int_0^t \frac{\hat{S}^w(s-)}{S(s)}d\hat{\Lambda}(s) \\ &= 1 - \int_0^t \frac{\hat{S}^w(s-)}{S(s)}\left\{\frac{dN^w_{\cdot}(s)}{Y^w_{\cdot}(s)} - d\Lambda(s)\right\}. \end{aligned}$$

This can be used to obtain the difference between the weighted estimator and the survival function as $\hat{S}^w(t) - S(t) = -S(t)\int_0^t \frac{\hat{S}^w(s-)}{S(s)}\left\{\frac{dN^w_{\cdot}(s)}{Y^w_{\cdot}(s)} - d\Lambda(s)\right\}$. Plugging this into (2.3.3), we can write $\sqrt{n}(\hat{F}^w_1(t) - F_1(t))$

$$\begin{aligned} &= \sqrt{n}\int_0^t \hat{S}^w(s-)\frac{1}{Y^w_{\cdot}(s)}\{dN^w_{1\cdot}(s) - Y^w_{\cdot}(s)d\Lambda_1(s)\} \\ &\quad -\sqrt{n}\int_0^t S(s)\left\{\int_0^s \frac{\hat{S}^w(u-)}{S(u)}\frac{1}{Y^w_{\cdot}(u)}dM^w_{\cdot}(u)\right\}d\Lambda_1(s) \\ &= \sqrt{n}\int_0^t \hat{S}^w(s-)\frac{1}{\sum_{i=1}^n Q_i}\frac{1}{\frac{Y^w_{\cdot}(s)}{\sum_{i=1}^n Q_i}}\{dN^w_{1\cdot}(s) - Y^w_{\cdot}(s)d\Lambda_1(s)\} \\ &\quad -\sqrt{n}\int_0^t S(s)\left\{\int_0^s \frac{\hat{S}^w(u-)}{S(u)}\frac{1}{\sum_{i=1}^n Q_i}\frac{1}{\frac{Y^w_{\cdot}(u)}{\sum_{i=1}^n Q_i}}dM^w_{\cdot}(u)\right\}d\Lambda_1(s) \\ &= \frac{1}{\sqrt{n}}\int_0^t \frac{n}{\sum_{i=1}^n Q_i}\hat{S}^w(s-)\frac{1}{\bar{Y}^w(s)}[dN^w_{1\cdot}(s) - Y^w_{\cdot}(s)d\Lambda_1(s)] \\ &\quad -\frac{1}{\sqrt{n}}\int_0^t \frac{n}{\sum_{i=1}^n Q_i}S(s)\left\{\int_0^s \frac{\hat{S}^w(u-)}{S(u)}\frac{1}{\bar{Y}^w(s)}dM^w_{\cdot}(u)\right\}d\Lambda_1(s), \end{aligned} \quad (2.3.4)$$

where $\bar{Y}^w(s) = \frac{Y^w(s)}{\sum_{i=1}^n Q_i}$. Using Taylor's expansion on $\frac{1}{\bar{Y}^w(s)}$, we can replace it with $\frac{1}{S(s-)}$ plus a smaller order term. The first term in (2.3.4) can be written as

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t \hat{S}^w(s-)\frac{1}{S(s-)}[dN_{1j}^w(s) - Y_j^w(s)d\Lambda_1(s)] + o_p(1) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} I_{1j}(t) + o_p(1).$$

Following a similar approach the second term in (2.3.4) can be written as

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t S(s)\left[\int_0^s \frac{\hat{S}^w(u-)}{S(u)}\frac{1}{S(u-)}(dN_j^w(u) - Y_j^w(u)d\Lambda(u))\right]d\Lambda_1(s) + o_p(1)\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} I_{2j}(t) + o_p(1).\end{aligned}$$

Thus, $\sqrt{n}(\hat{F}_1^w(t) - F_1(t)) = \frac{1}{\sqrt{n}}\sum_{j=1}^n I_j(t) + o_p(1)$ where $I_j(t) = I_{1j}(t) - I_{2j}(t)$. As a result we obtain the following variance estimator for the weighted CIF estimator:

$$\hat{\sigma}^2(\hat{F}_1^w(t)) = \frac{1}{n^2}\sum_{j=1}^{n}\hat{I}_j^2(t),$$

where $\hat{I}_j(t) = \hat{I}_{1j}(t) - \hat{I}_{2j}(t)$ with

$$\begin{aligned}\hat{I}_{1j}(t) &= \int_0^t dN_{1j}^w(s) - Y_j^w(s)d\Lambda_1(s)\\ &= Q_j\sum_{t_m\le t}\left[I(V_j = t_m, \Delta_j = 1) - I(V_j \ge t_m)\frac{d_m^w}{Y_m^w}\right]\end{aligned}$$

and

$$\begin{aligned}\hat{I}_{2j}(t) &= \int_0^t \hat{S}^w(s)\left[\int_0^s \frac{1}{\hat{S}^w(u)}(dN_j^w(u) - Y_j^w(u)d\Lambda(u))\right]d\Lambda_1(s)\\ &= Q_j\sum_{t_m\le t}\hat{S}^w(t_m)\left[\sum_{t_d\le t_m}\frac{1}{\hat{S}^w(t_d)}\left(I(V_j = t_d) - I(V_j \ge t_d)\frac{d_d^w + r_d^w}{Y_d^w}\right)\right]\frac{d_m^w}{Y_m^w}.\end{aligned}$$

2.3.2 The CIF Estimator with Estimated Fixed Weights

In practice, due to randomization, the proportion of subjects who responded to the initial treatment A_1 and were randomized to B_k may not be exactly the same as π_{B_k}. Similarly, the proportion of subjects receiving B_l' could be different from $\pi_{B_l'}$. In some situations it may be necessary to estimate the randomization probabilities for the second stage or the sample proportion may provide better information about the randomization process than the intended assignment probabilities. For these reasons, we also propose a weighted CIF estimator where the weights are estimated using the sample proportions instead of true probabilities:

$$\hat{F}^{ew}_{1,A_1B_kB_l'}(t) = \sum_{t_i \le t} \frac{d_i^{ew}}{Y_i^{ew}} \left\{ \prod_{j=1}^{i-1} \left(1 - \frac{d_j^{ew} + r_j^{ew}}{Y_j^{ew}} \right) \right\}, \tag{2.3.5}$$

for $t_1 \le t$ and 0 otherwise, where $d_i^{ew} = \sum_{j=1}^n I(\Delta_j = 1) I(V_j = t_i) \hat{Q}_{A_1B_kB_l',j}$, $r_i^{ew} = \sum_{j=1}^n I(\Delta_j = 2) I(V_j = t_i) \hat{Q}_{A_1B_kB_l',j}$ and $Y_i^{ew} = \sum_{j=1}^n I(V_j \ge t_i) \hat{Q}_{A_1B_kB_{lj}'}$ for $\hat{Q}_{A_1B_kB_l',j} = R_j I\{Z_{1j} = k\}/\hat{\pi}_{B_k} + (1 - R_j) I\{Z_{2j} = l\}/\hat{\pi}_{B_l'}$. The variance of this estimator can be estimated by replacing the weights with their estimated values in the formula derived for the CI with fixed weights.

2.3.3 The CIF Estimators with Time-Dependent Weights

The proposed cumulative incidence functions with fixed weights and estimated fixed weights can be improved in a way so that more subjects provide information for the estimation of the cumulative incidence for a given regime. To do this, following the ideas from Guo and Tsiatis (2005), subjects can be given weights of 1 until their response status are observed because they remain consistent with all of the regimes. Once the response status is known and the second randomization is carried out, the patients receive weights according to the regimes they follow. The weights evaluated at time t can be written as below:

$$Q_{A_1B_kB_l',j}(t) = \begin{cases} 1, & \text{if } T_j^R > t, \\ \frac{R_j I\{Z_{1j}=k\}}{\pi_{B_k}} + \frac{(1-R_j) I\{Z_{2j}=l\}}{\pi_{B_l'}}, & \text{if } T_j^R \le t. \end{cases}$$

Using the time-dependent weights the cumulative incidence for a specific regime can be estimated as

$$\hat{F}^{tw}_{A_1B_kB_l'}(t) = \sum_{t_i \le t} \frac{d_i^{tw}}{Y_i^{tw}} \left\{ \prod_{j=1}^{i-1} \left(1 - \frac{d_j^{tw} + r_j^{tw}}{Y_j^{tw}} \right) \right\}. \tag{2.3.6}$$

Here $\hat{F}^{tw}_{A_1B_kB_l'}$ denotes the estimated CIF with time-dependent weights and

$$\begin{aligned} d_i^{tw} &= \sum_{j=1}^{n} I(\Delta_j = 1) I(V_j = t_i) Q_{A_1B_kB_l',j}(t_i), \\ r_i^{tw} &= \sum_{j=1}^{n} I(\Delta_j = 2) I(V_j = t_i) Q_{A_1B_kB_l',j}(t_i), \\ Y_i^{tw} &= \sum_{j=1}^{n} I(V_j \ge t_i) Q_{A_1B_kB_l',j}(t_i). \end{aligned}$$

The associated influence function can be obtained by a slight modification of the previous influence function and the new variance estimator can be obtained just by replacing the two parts of the influence function with the ones below:

$$\begin{aligned} \hat{I}_{1j}^{tw}(t) &= Q_j(t) \sum_{t_m \le t} \left[I(V_j = t_m, \Delta_j = 1) - I(V_j \ge t_m) \frac{d_m^{tw}}{Y_m^{tw}} \right] \\ \hat{I}_{2j}^{tw}(t) &= Q_j(t) \sum_{t_m \le t} \hat{S}^{tw}(t_m) \left[\sum_{t_d \le t_m} \frac{1}{\hat{S}^{tw}(t_d)} \left(I(V_j = t_d) - I(V_j \ge t_d) \frac{d_d^{tw} + r_d^{tw}}{Y_d^{tw}} \right) \right] \frac{d_m^{tw}}{Y_m^{tw}}. \end{aligned}$$

The CIF estimator with estimated time-dependent weights $\hat{F}^{tew}_{A_1B_kB_l'}$ and its variance estimator can be obtained by replacing the weights in $\hat{F}^{tw}_{A_1B_kB_l'}$ and its variance with the estimated ones.

2.4 COMPARING TWO REGIMES

2.4.1 Confidence Intervals and Confidence Bands

Suppose we are interested in comparing the two regimes $A_1B_1B_1'$ and $A_1B_1B_2'$ which share some common path. We would then be interested in the difference $D(t) = F_{1,A_1B_1B_1'}(t) - F_{1,A_1B_1B_2'}(t)$. It can be consistently estimated by the difference between the two estimated CIFs with time-dependent weights of the respective regimes, i.e., $\hat{D}^{tw}(t) = \hat{F}^{tw}_{1,A_1B_1B_1'}(t) - \hat{F}^{tw}_{1,A_1B_1B_2'}(t)$. If we denote the influence function for the two regimes respectively as $I_j^{(1)}(t)$ and $I_j^{(2)}(t)$, we can write:

$$\begin{aligned}\sqrt{n}(\hat{D}^{tw}(t) - D(t)) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{I_j^{(1)}(t) - I_j^{(2)}(t)\right\} + o_p(1)\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} I_j^{D}(t) + o_p(1).\end{aligned}$$

Since $I_j^D(t)$ can be easily estimated by $\hat{I}_j^D(t) = \hat{I}_j^{(1)}(t) - \hat{I}_j^{(2)}(t)$ and the variance of $\hat{D}^{tw}$ can be estimated as $\hat{\sigma}^2_{\hat{D}(t)} = \frac{1}{n^2}\sum_{j=1}^{n}\left\{\hat{I}_j^D(t)\right\}^2$, the $100(1-\alpha)\%$ confidence interval for $D(t)$ is

$$\left\{\hat{F}^{tw}_{1,A_1B_1B_1'}(t) - \hat{F}^{tw}_{1,A_1B_1B_2'}(t)\right\} \pm Z_{\alpha/2}\hat{\sigma}_{\hat{D}(t)},$$

where $P(N(0,1) \geq Z_{\alpha/2}) = \alpha/2$.

In addition to point-wise confidence intervals, confidence bands are often constructed for functions of $D(t)$ to determine the time regions where the two CIFs differ. We adapt Lin's re-sampling technique (Lin et al., 1994; Lin, 1997) following the guidelines of Zhang and Fine (2008). More specifically, we consider a general transformation $G(F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t))$. Let $G^{(1)}(u,v) = \partial G(u,v)/\partial u$ and $G^{(2)}(u,v) = \partial G(u,v)/\partial v$ be first-order partial derivatives of G, and n_1 and n_2 be the numbers of subjects who are consistent with the regimes $A_1B_1B_1'$ and $A_1B_1B_2'$. We assume that $n_i/(n_1+n_2) \to \rho_i$ as $n_1+n_2 \to \infty$, where $0 < \rho_i < 1, i = 1,2$.

Then the functional delta method gives that

$$\begin{aligned}
& \sqrt{n}[G(\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t)) - G(F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t))] \\
= & \sum_{i=1}^{2}\sum_{j=1}^{n_i} \rho_i^{-1/2} G^{(i)}(F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t)) n_i^{-1/2} I_j^{(i)}(t) + o_P(1) \\
= & \sqrt{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i} I_j^{(i)G} + o_P(1),
\end{aligned}$$

where $I_j^{(i)G}$ can be estimated by $\hat{I}_j^{(i)G} = G^{(i)}(\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t))\hat{I}_j^{(i)}(t)/n_i$.

Thus, $\sqrt{n}[G(\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t)) - G(F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t)])$ converges weakly to a Gaussian process with a variance consistently estimated by:

$$\hat{\Sigma}_G(t) = n\sum_{i=1}^{2}\sum_{j=1}^{n_i}\{\hat{I}_j^{(i)G}(t)\}^2.$$

Let $Z_{ij}^{(b)}, i = 1, 2, j = 1, \ldots, n_i, b = 1, \ldots, N$ be independent standard normal variates. Then the simulated process

$$\hat{J}^{(b)}(t) = \sqrt{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i} G^{(i)}(\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t))\hat{I}_j^{(i)}(t)Z_{ij}^{(b)}/n_i$$

has the same limiting process as:

$$\sqrt{n}[G(\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t)) - G(F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t))].$$

Let C_α be $100(1-\alpha)^{th}$ percentile of

$$\bar{J}^{(b)} = \sup_{t\epsilon[\tau_l,\tau_u]} |\hat{J}^{(b)}(t)|,$$

for $b = 1, \ldots, N$ and $[\tau_l, \tau_u] \subset [0, \tau]$. Then $100(1-\alpha)\%$ confidence bands for the transformation $G(F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t))$ are:

$$G(\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t)) \pm C_\alpha\sqrt{\hat{\Sigma}_G(t)/n}. \tag{2.4.1}$$

In the simple case where $G(u, v) = u - v$ the simulated process $\hat{J}^{(b)}$ can be written as:

$$\hat{J}^{(b)} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{\hat{I}_j^{(1)}Z_{1j}^{(b)} - \hat{I}_j^{(2)}Z_{2j}^{(b)}\right\}$$

and the variance estimator is $\hat{\Sigma}_G(t) = \frac{1}{n}\sum_{j=1}^{n}\left(\hat{I}_j^{(1)} - \hat{I}_j^{(2)}\right)^2$. Now, the confidence bands can be computed as in (2.4.1).

2.4.2 Time-averaged Differences

In practice, one is often interested in summarizing the difference between two CIFs over time to obtain a global measure of the difference between two treatment regimes. Let G be some general distance measure. To combine information in $G\{F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t)\}$ over time, Zhang and Fine (2008) proposed weighted average summaries:

$$G_M = \int_{\tau_l}^{\tau_u} G\{F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t)\} dW(t),$$

where $W(t) > 0$ is a deterministic weight function and $\int_{\tau_l}^{\tau_u} dW(t) = 1$. Following the ideas of the weighted log-rank tests for censored survival data, one may consider the class of weights based on the CIF calculated by pooling the data from both regimes. The estimator for the time-averaged difference is:

$$\hat{G}_M = \int_{\tau_l}^{\tau_u} G\{\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t)\} dW(t).$$

Then $\sqrt{n}(\hat{G}_M - G_M)$ can be expressed as

$$\begin{aligned} &= \sqrt{n} \int_{\tau_l}^{\tau_u} [G\{\hat{F}^{tw}_{1,A_1B_1B_1'}(t), \hat{F}^{tw}_{1,A_1B_1B_2'}(t)\} - G\{F_{1,A_1B_1B_1'}(t), F_{1,A_1B_1B_2'}(t)\}] dW(t) \\ &= \sqrt{n} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \hat{I}_j^{(i)G_M} + o_p(1), \end{aligned}$$

where $\hat{I}_j^{(i)G_M} = \int_{\tau_l}^{\tau_u} \hat{I}_j^{(i)G}(t) dW(t)$. The asymptotic variance can then be estimated by $\hat{\Sigma}_{G_M} = n \sum_{i=1}^{2} \sum_{j=1}^{n_i} \{\hat{I}_j^{(i)G_M}\}^2$. For the case where $G(u, v) = u - v$, $\hat{I}_j^{(1)G}(t) = \hat{I}_j^{(1)}/n$ and $\hat{I}_j^{(2)G}(t) = -\hat{I}_j^{(2)}/n$. Inference can easily be carried out based on these influence functions.

2.5 SIMULATION

Simulation studies were carried out to compare the proposed estimator to the standard estimator under different conditions. In all simulations the study design explained in Figure 1 was used. Only the subjects who received the initial treatment A_1 were considered since the data obtained from these subjects is independent of the data obtained from the subjects that were assigned to the other initial treatment A_2. The comparisons were made under three different scenarios. Namely, the data were generated at which 1) $Z_{1i} \sim 2 - Bernoulli(0.5)$ and $Z_{2i} \sim 2 - Bernoulli(0.5)$, 2) $Z_{1i} \sim 2 - Bernoulli(0.3)$ and $Z_{2i} \sim 2 - Bernoulli(0.3)$ (i.e. $Z_{1i} = 1$ with probability 0.3 and $Z_{1i} = 2$ with probability 0.7), and 3) $Z_{1i} = 1$ and $Z_{2i} \sim 2 - Bernoulli(0.5)$. Every model was repeated for $n = 300, 700$ and for cases $R_i \sim Bernoulli(0.4)$ and $R_i \sim Bernoulli(0.7)$ and 4000 data sets were generated for each setting.

For each combination, $\{(T_i^R, R_i, Z_{1i}, Z_{2i}, V_i, \Delta_i), i = 1, \ldots, n\}$ were generated. More specifically, T_i^R, the times to response, were generated from Exponential (0.20) and restricted at 1 year. The times to death from the second randomization ($T^*_{A_1B_{ki}}$ or $T^*_{A_1B'_{li}}$, $k, l = 1, 2$) were drawn from different exponential distributions with the parameter values of 1 for the sequence of treatments A_1B_1, 0.75 for the A_1B_2, 0.50 for the A_1B_1' and 0.25 for the A_1B_2' treatments. Following Miyahara and Wahed (2010), we then defined the overall survival time for subject i as $T_i = T_i^R + R_i\{I(Z_{1i} = 1)T^*_{A_1B_{1i}} + I(Z_{1i} = 2)T^*_{A_1B_{2i}}\} + (1 - R_i)\{I(Z_{2i} = 1)T^*_{A_1B'_{1i}} + I(Z_{2i} = 2)T^*_{A_1B'_{2i}}\}$. The times to censoring C_i were generated from a Uniform $(1.5, 2)$ which resulted in 9% censoring for $P(R = 1) = 0.4$ and 13% censoring for $P(R = 1) = 0.7$. Only the results for the regimes $A_1B_1B_1'$ and $A_1B_1B_2'$ were given since the results for other regimes were similar.

In the following tables, CI(t) stands for the naïve estimate of the CIF in (2.3.1) evaluated at time t, WCI(t) stands for the proposed estimate of the CIF with fixed weights in (2.3.2) and WCI2(t) is the proposed weighted CIF estimate with estimated fixed weights in (2.3.5), TWCI(t) is the estimate of the CIF with time-dependent weights in (2.3.6) and TWCI2(t) is the CIF estimate with estimated time-dependent weights.

It is worth noting that although one might think of the simulated sample sizes as large, because only the data from a specific regime were used in the estimation process, the sample used in the estimation of the CIF for a particular regime is significantly smaller than the entire sample.

Table 1 shows the simulation results for Scenario-1. It can be seen that all methods produce comparable results to the standard estimator under both response rates. The coverage rates for all estimators improve and get closer to the desired value 95% as n gets larger. The results for the two regimes are similar. This is not surprising, as when $\pi_{B_1} = \pi_{B_1'} = 0.5$, the responders receiving A_1B_1 and non-responders receiving A_1B_1' are assigned approximately equal weights in all the weighting schemes. The pseudo sample after weighting has roughly the same mixture of responders and non-responders as the original sample. Therefore, the weighted methods produce similar estimates to the naïve estimator.

The simulation results for the second scenario can be seen in Table 2. Under the second scenario, for the first regime $A_1B_1B_1'$ the naïve and weighted methods perform similarly due to the fact that $\pi_{B_1} = \pi_{B_1'} = 0.3$. This is again not surprising, as when $\pi_{B_1} = \pi_{B_1'} = 0.3$, the responders receiving A_1B_1 and non-responders receiving A_1B_1' are assigned approximately equal weight. As a result, the pseudo sample created for the first regime after weighting has roughly the same mixture of responders and non-responders as the original sample. Therefore, the weighted methods produce similar estimates to the naïve estimator. However, for the second regime $A_1B_1B_2'$, the naïve CIF estimator produces biased results. The naïve estimate is obtained based on the data from about 30% of responders who actually received A_1B_1 and about 70% of non-responders who actually received A_1B_2'. However, the remaining responders and non-responders could be equally qualified for receiving this treatment regime if there were no second-stage randomization. The weighted methods roughly generate a pseudo sample that represents all responders and non-responders, which in turn produce more accurate estimates of the true CIF than the naïve estimator. The coverage rates from the proposed weighted methods improve as the sample size is increased unlike the standard estimator. TWCI2 performs slightly better compared to the other proposed methods.

The results from the third scenario can be seen in Table 3. It can be clearly seen that

Table 1: Inverse weighting- Scenario-1: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation (σ), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

				$A_1B_1B_1'$					$A_1B_1B_2'$				
n	t	pr	Method	True	Mean	σ	$\hat{\sigma}$	Cov	True	Mean	σ	$\hat{\sigma}$	Cov
300	0.5	0.4	CI	0.185	0.185	0.032	0.032	0.941	0.248	0.248	0.036	0.036	0.946
			WCI		0.185	0.032	0.031	0.938		0.248	0.036	0.035	0.940
			WCI2		0.185	0.032	0.031	0.939		0.248	0.036	0.035	0.940
			TWCI		0.185	0.031	0.031	0.939		0.248	0.035	0.035	0.940
			TWCI2		0.185	0.031	0.031	0.941		0.248	0.035	0.035	0.940
		0.7	CI	0.158	0.158	0.030	0.030	0.942	0.190	0.189	0.032	0.032	0.946
			WCI		0.158	0.030	0.029	0.940		0.189	0.032	0.032	0.942
			WCI2		0.158	0.030	0.029	0.939		0.189	0.032	0.031	0.943
			TWCI		0.158	0.030	0.029	0.940		0.189	0.032	0.031	0.943
			TWCI2		0.158	0.030	0.029	0.941		0.189	0.032	0.031	0.943
	1	0.4	CI	0.342	0.342	0.038	0.041	0.960	0.389	0.389	0.041	0.042	0.961
			WCI		0.342	0.038	0.038	0.948		0.389	0.041	0.039	0.942
			WCI2		0.342	0.038	0.038	0.950		0.389	0.040	0.039	0.942
			TWCI		0.342	0.038	0.039	0.950		0.389	0.040	0.040	0.946
			TWCI2		0.342	0.038	0.039	0.951		0.389	0.040	0.040	0.947
		0.7	CI	0.306	0.307	0.038	0.039	0.956	0.330	0.330	0.038	0.040	0.961
			WCI		0.307	0.038	0.037	0.948		0.330	0.038	0.038	0.952
			WCI2		0.307	0.038	0.037	0.947		0.330	0.037	0.038	0.953
			TWCI		0.307	0.038	0.037	0.950		0.330	0.038	0.038	0.953
			TWCI2		0.307	0.038	0.037	0.948		0.330	0.038	0.038	0.956
700	0.5	0.4	CI	0.185	0.185	0.031	0.032	0.946	0.248	0.248	0.035	0.036	0.949
			WCI		0.185	0.031	0.031	0.944		0.248	0.035	0.035	0.942
			WCI2		0.185	0.031	0.031	0.944		0.248	0.035	0.035	0.939
			TWCI		0.185	0.031	0.031	0.944		0.248	0.035	0.035	0.941
			TWCI2		0.185	0.031	0.031	0.946		0.248	0.035	0.035	0.942
		0.7	CI	0.157	0.157	0.030	0.030	0.940	0.190	0.189	0.032	0.032	0.942
			WCI		0.157	0.030	0.029	0.936		0.189	0.032	0.032	0.939
			WCI2		0.157	0.030	0.029	0.935		0.189	0.032	0.032	0.940
			TWCI		0.157	0.030	0.029	0.936		0.189	0.032	0.031	0.941
			TWCI2		0.157	0.030	0.029	0.938		0.189	0.032	0.031	0.938
	1	0.4	CI	0.342	0.342	0.038	0.041	0.964	0.389	0.389	0.040	0.043	0.962
			WCI		0.342	0.038	0.038	0.952		0.389	0.040	0.039	0.944
			WCI2		0.342	0.038	0.038	0.952		0.389	0.039	0.039	0.947
			TWCI		0.342	0.038	0.039	0.956		0.389	0.040	0.040	0.949
			TWCI2		0.342	0.038	0.039	0.955		0.389	0.039	0.040	0.952
		0.7	CI	0.306	0.306	0.037	0.039	0.958	0.330	0.330	0.038	0.040	0.963
			WCI		0.306	0.037	0.037	0.948		0.330	0.038	0.038	0.948
			WCI2		0.306	0.037	0.037	0.949		0.330	0.038	0.038	0.948
			TWCI		0.306	0.037	0.037	0.951		0.330	0.038	0.038	0.951
			TWCI2		0.306	0.037	0.037	0.952		0.330	0.038	0.038	0.952

Table 2: Inverse weighting- Scenario-2: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation (σ), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

				$A_1B_1B_1'$					$A_1B_1B_2'$				
n	t	pr	Method	True	Mean	σ	$\hat{\sigma}$	Cov	True	Mean	σ	$\hat{\sigma}$	Cov
300	0.5	0.4	CI	0.185	0.185	0.041	0.041	0.939	0.248	0.283	0.036	0.036	0.863
			WCI		0.185	0.041	0.040	0.934		0.248	0.035	0.034	0.937
			WCI2		0.185	0.041	0.040	0.934		0.248	0.035	0.034	0.940
			TWCI		0.185	0.041	0.040	0.940		0.248	0.035	0.034	0.942
			TWCI2		0.185	0.041	0.040	0.938		0.248	0.035	0.034	0.942
		0.7	CI	0.158	0.158	0.038	0.038	0.943	0.190	0.229	0.038	0.038	0.846
			WCI		0.158	0.038	0.038	0.938		0.190	0.036	0.035	0.935
			WCI2		0.158	0.038	0.038	0.939		0.190	0.036	0.035	0.937
			TWCI		0.158	0.038	0.037	0.940		0.190	0.035	0.034	0.936
			TWCI2		0.158	0.038	0.038	0.940		0.190	0.035	0.035	0.936
	1	0.4	CI	0.342	0.343	0.051	0.052	0.950	0.389	0.425	0.039	0.042	0.891
			WCI		0.343	0.051	0.049	0.938		0.390	0.041	0.040	0.944
			WCI2		0.343	0.051	0.049	0.941		0.390	0.041	0.040	0.946
			TWCI		0.343	0.051	0.050	0.942		0.390	0.041	0.041	0.954
			TWCI2		0.343	0.051	0.050	0.946		0.390	0.040	0.041	0.954
		0.7	CI	0.306	0.307	0.049	0.050	0.947	0.330	0.370	0.043	0.043	0.882
			WCI		0.307	0.049	0.048	0.934		0.331	0.045	0.043	0.938
			WCI2		0.307	0.049	0.048	0.937		0.331	0.044	0.043	0.939
			TWCI		0.307	0.049	0.048	0.940		0.331	0.044	0.044	0.943
			TWCI2		0.307	0.049	0.048	0.942		0.330	0.044	0.044	0.943
700	0.5	0.4	CI	0.185	0.186	0.042	0.041	0.936	0.248	0.283	0.035	0.036	0.865
			WCI		0.186	0.042	0.040	0.929		0.249	0.035	0.034	0.945
			WCI2		0.186	0.042	0.040	0.930		0.249	0.034	0.034	0.948
			TWCI		0.186	0.041	0.040	0.932		0.249	0.034	0.034	0.947
			TWCI2		0.186	0.041	0.040	0.932		0.249	0.034	0.034	0.947
		0.7	CI	0.158	0.158	0.039	0.038	0.931	0.190	0.228	0.037	0.038	0.858
			WCI		0.158	0.039	0.038	0.924		0.190	0.035	0.035	0.938
			WCI2		0.158	0.039	0.038	0.926		0.190	0.035	0.035	0.940
			TWCI		0.158	0.039	0.037	0.932		0.189	0.035	0.034	0.938
			TWCI2		0.158	0.039	0.037	0.930		0.189	0.035	0.035	0.938
	1	0.4	CI	0.342	0.343	0.051	0.052	0.954	0.389	0.425	0.039	0.042	0.902
			WCI		0.343	0.051	0.049	0.941		0.390	0.041	0.040	0.942
			WCI2		0.343	0.051	0.049	0.940		0.390	0.041	0.040	0.944
			TWCI		0.342	0.051	0.050	0.945		0.390	0.041	0.041	0.949
			TWCI2		0.343	0.050	0.050	0.943		0.390	0.040	0.041	0.951
		0.7	CI	0.306	0.306	0.049	0.050	0.954	0.330	0.369	0.043	0.045	0.898
			WCI		0.306	0.049	0.048	0.942		0.330	0.044	0.043	0.938
			WCI2		0.306	0.049	0.048	0.944		0.330	0.044	0.043	0.937
			TWCI		0.306	0.049	0.048	0.945		0.330	0.043	0.044	0.943
			TWCI2		0.306	0.049	0.048	0.948		0.329	0.043	0.044	0.944

Table 3: Inverse weighting- Scenario-3: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation (σ), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

				$A_1B_1B_1'$					$A_1B_1B_2'$				
n	t	pr	Method	True	Mean	σ	$\hat{\sigma}$	Cov	True	Mean	σ	$\hat{\sigma}$	Cov
300	0.5	0.4	CI	0.185	0.170	0.026	0.026	0.883	0.248	0.215	0.028	0.029	0.779
			WCI		0.185	0.029	0.029	0.939		0.248	0.032	0.032	0.946
			WCI2		0.185	0.029	0.029	0.940		0.249	0.032	0.032	0.949
			TWCI		0.185	0.029	0.029	0.939		0.248	0.032	0.032	0.942
			TWCI2		0.185	0.029	0.029	0.940		0.248	0.032	0.032	0.948
		0.7	CI	0.158	0.147	0.022	0.022	0.905	0.190	0.166	0.023	0.023	0.811
			WCI		0.158	0.025	0.025	0.935		0.190	0.027	0.027	0.939
			WCI2		0.158	0.025	0.025	0.936		0.190	0.027	0.027	0.939
			TWCI		0.158	0.025	0.025	0.938		0.190	0.027	0.027	0.939
			TWCI2		0.158	0.025	0.025	0.938		0.190	0.027	0.027	0.941
	1	0.4	CI	0.342	0.322	0.032	0.034	0.905	0.389	0.356	0.033	0.035	0.838
			WCI		0.342	0.035	0.035	0.946		0.389	0.036	0.036	0.946
			WCI2		0.342	0.035	0.035	0.946		0.389	0.036	0.036	0.948
			TWCI		0.342	0.035	0.035	0.946		0.389	0.036	0.036	0.946
			TWCI2		0.342	0.035	0.035	0.947		0.389	0.036	0.036	0.949
		0.7	CI	0.306	0.292	0.028	0.029	0.924	0.330	0.306	0.028	0.030	0.881
			WCI		0.307	0.031	0.031	0.945		0.330	0.031	0.031	0.950
			WCI2		0.307	0.031	0.031	0.944		0.330	0.031	0.031	0.949
			TWCI		0.307	0.031	0.031	0.945		0.330	0.031	0.031	0.948
			TWCI2		0.307	0.031	0.031	0.946		0.330	0.031	0.031	0.950
700	0.5	0.4	CI	0.185	0.170	0.026	0.026	0.886	0.248	0.215	0.028	0.029	0.772
			WCI		0.185	0.029	0.029	0.938		0.248	0.033	0.032	0.943
			WCI2		0.185	0.029	0.029	0.940		0.248	0.033	0.032	0.944
			TWCI		0.185	0.029	0.029	0.938		0.248	0.033	0.032	0.942
			TWCI2		0.185	0.029	0.029	0.938		0.248	0.032	0.032	0.946
		0.7	CI	0.158	0.147	0.022	0.022	0.899	0.190	0.166	0.023	0.023	0.802
			WCI		0.158	0.025	0.025	0.940		0.190	0.028	0.027	0.939
			WCI2		0.158	0.025	0.025	0.942		0.190	0.027	0.027	0.940
			TWCI		0.158	0.025	0.025	0.940		0.190	0.028	0.027	0.936
			TWCI2		0.158	0.025	0.025	0.942		0.190	0.027	0.027	0.940
	1	0.4	CI	0.342	0.322	0.033	0.034	0.902	0.389	0.356	0.034	0.035	0.836
			WCI		0.342	0.036	0.035	0.945		0.389	0.037	0.036	0.940
			WCI2		0.342	0.036	0.035	0.945		0.389	0.036	0.036	0.941
			TWCI		0.342	0.036	0.035	0.946		0.389	0.037	0.036	0.942
			TWCI2		0.342	0.036	0.035	0.945		0.389	0.036	0.036	0.944
		0.7	CI	0.306	0.292	0.029	0.029	0.914	0.330	0.306	0.029	0.030	0.874
			WCI		0.307	0.031	0.031	0.942		0.331	0.032	0.031	0.942
			WCI2		0.307	0.031	0.031	0.942		0.331	0.032	0.031	0.944
			TWCI		0.307	0.032	0.031	0.943		0.331	0.032	0.031	0.941
			TWCI2		0.307	0.031	0.031	0.945		0.331	0.032	0.031	0.942

the naïve estimator consistently produces biased estimates for both regimes. In contrast, all proposed weighted methods are unbiased for both response rates and both regimes. The coverage rates for the proposed methods are much better especially for the second regime where the naïve estimator coverage rate could go as low as 0.772 compared to TWCI2 achieving 0.946 under the same conditions.

2.6 DATA ANALYSIS

The proposed methods were applied to the data set from the Cancer and Leukemia Group B (CALGB) trial which was a two-stage randomized trial conducted to evaluate the effects of adding granulocyte-macrophage colony-stimulating factor (GM-CSF) to standard chemotherapy on elderly patients with leukemia (Stone et al., 1995). A total of 388 elderly patients with acute myelogenous leukemia were enrolled and the study was double-blinded and placebo-controlled. Upon entrance to the study 193 patients were randomized to the initial treatment where they received GM-CSF (A_1) in addition to the standard chemotherapy and 195 were randomized to the Chemo (A_2) group where they received placebo in addition to the standard chemotherapy. Response for this trial was defined as complete remission. Responders were randomized so that at the second stage 37 GM-CSF and 45 Chemo responders received the maintenance therapy 1 (B_1) and 42 GM-CSF and 45 Chemo responders were randomized to the maintenance therapy 2 (B_2). The non-responders were not randomized at the second stage. Therefore, there are a total of four regimes in this study A_1B_1, A_1B_2, A_2B_1 and A_2B_2, where $A_kB_l, k, l = 1, 2$ denotes the treatment policy that a subject started with A_k, and then received B_l if the subject responded and no further treatment if the subject did not respond. Here, the fixed weight function for the regime A_1B_k can be written as

$$Q_{A_1B_{ki}} = \frac{R_i I(Z_{1i} = k)}{\pi_{B_k}} + (1 - R_i), k = 1, 2.$$

Here $\pi_{B_1} = \pi_{B_2} = 0.5$ and the time-dependent weight function can be written as

$$Q_{A_1 B_{ki}}(t) = \begin{cases} 1, & \text{if } I(T_i^R > t) = 1, \\ \frac{R_i I(Z_{1i}=k)}{\pi_{B_k}} + (1 - R_i), & \text{if } I(T_i^R > t) = 0, \end{cases}$$

for $k = 1, 2$.

Although data were available to these authors for only one type of failure, for those who experienced an event and were not censored, we created a pseudo event type that was set equal to 1 or 2 with probability 0.5 for the purpose of demonstrating the proposed methods. We calculated the standard CIF estimator (CI), the CIF estimator with fixed weights (WCI) and the CIF estimator with time-dependent weights (TWCI) for the cause 1 event of these four regimes. The results are shown in Figure 2.

It can be seen from Figure 2 that the standard CIF estimator generally produces higher estimates for the CIF and the two weighted estimators produce very similar results. If there existed a type-1 event as we had simulated, we could then choose the regime that produces the lowest probability of occurrence for type-1 event by comparing the CIF values for the four regimes using the proposed weighted estimators. Since no such distinction exists in failure types in this data, to recommend a treatment regime, survival probabilities must be examined. We can obtain a weighted estimate for the overall survival function based on the estimated cause 1 and cause 2 CIFs. That is,

$$\hat{S}^w_{A_k B_l}(t) = 1 - \hat{F}^w_{1,A_k B_l}(t) - \hat{F}^w_{2,A_k B_l}(t),$$

for $k, l = 1, 2$. The survival probabilities for the four regimes calculated using the above formula are plotted in Figure 3. As it can be seen from this figure none of the regimes clearly out-perform the others, which are consistent with the results from the weighted survival estimates in Miyahara and Wahed (2010).

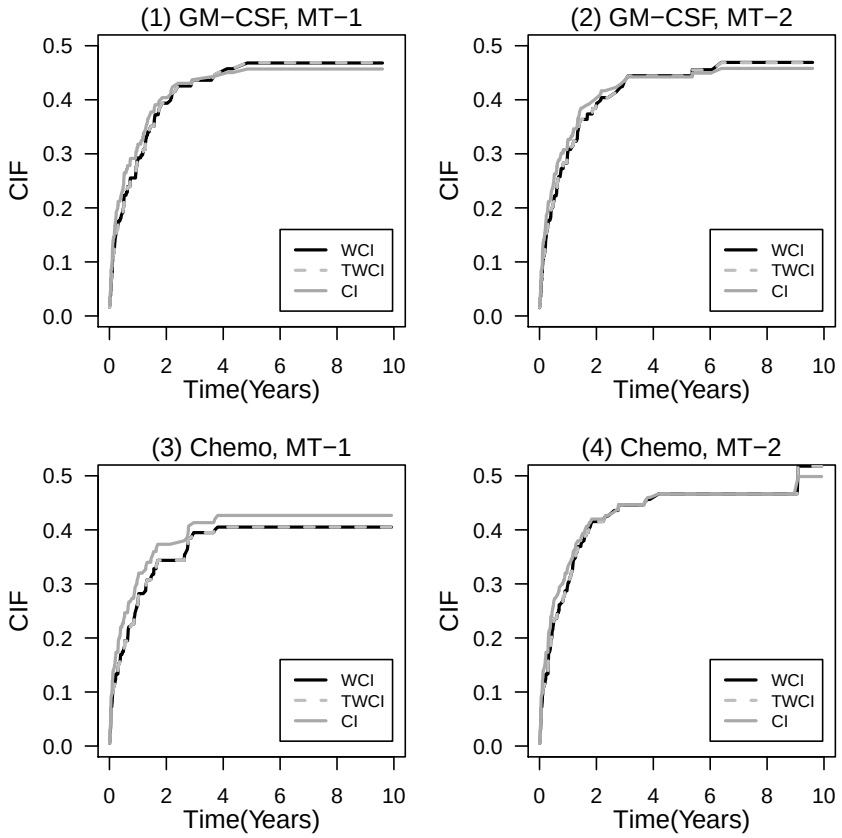

Figure 2: Inverse weighted and naïve CIF estimators for the four regimes.

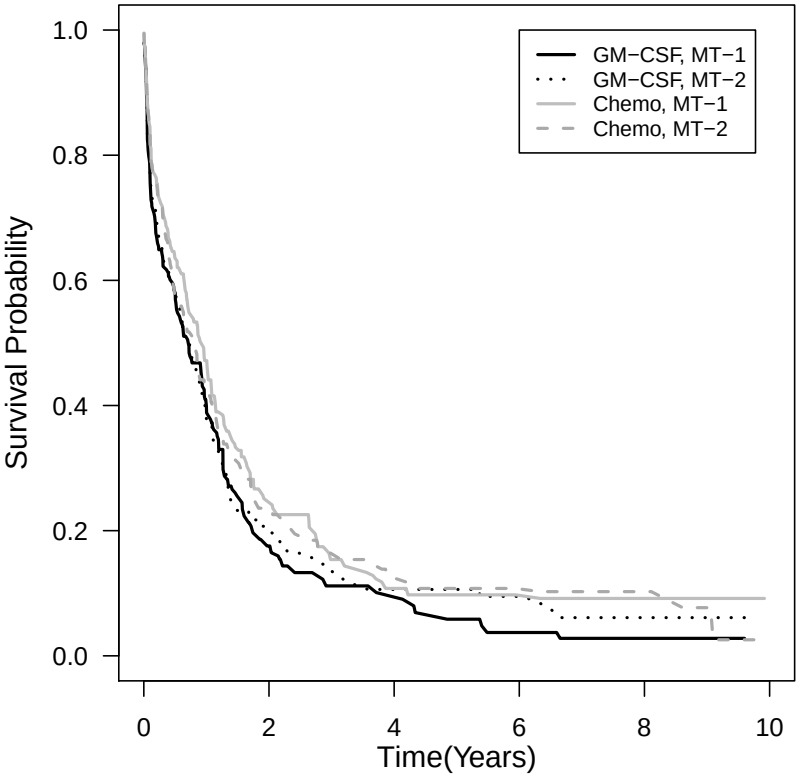

Figure 3: Comparison of regimes with respect to the weighted survival estimates.

2.7 SUMMARY

In this chapter we proposed and compared different inverse weighted estimators of the CIF which can be used when a two-stage randomization is carried out. The weighted methods are easy to implement with explicit variance estimators. Therefore, we were able to provide inference procedures to compare two treatment regimes and methods to obtain confidence intervals and bands. We showed that the proposed estimators produce unbiased estimates of the CIF with good coverage rates where the naïve estimator fails to do so. In general, the CIF estimator with estimated time-dependent weights produces slightly better results so it should be preferred when possible. Since dynamic treatment regimes are widely used in treating diseases that require complex treatment and competing-risk censoring is common in studies with multiple endpoints, the proposed methods provide useful inferential tools that will help advocate research in personalized medicine.

3.0 A PATTERN-MIXTURE TYPE ESTIMATOR FOR THE CUMULATIVE INCIDENCE FUNCTION UNDER TWO-STAGE RANDOMIZATION

3.1 INTRODUCTION

Pattern-mixture models (Little, 1993) are widely used in longitudinal data analysis when there are informative dropouts. The models yield a parameter estimate by first separating the data into different sections according to their dropout patterns, estimating the parameter on each section, and then combining these estimates as a weighted average. In causal inference literature this approach where the overall parameter estimate is obtained by first estimating the distribution of observed intermediate outcomes is referred to as the G-computation algorithm (Robins, 1986, 1987). This algorithm was borrowed by Thall et al. (2000) and Lavori and Dawson (2004) where they applied it to the analysis of data from sequentially randomized trials. Later Wahed (2010) developed inference procedures for mean survival times of two-stage adaptive treatment strategies using mixture distributions. Dawson and Lavori (2010) adapted the pattern-mixture models to sample size calculations for the purpose of evaluating treatment policies when a mean-response was of concern. Wahed and Tang (2013) utilized the pattern-mixture models to develop a Nelson-Aalen type estimator for survival. They proposed parametric models for the survival functions of treatment regimes, and tested for differences among these regimes.

In this chapter, we use similar ideas and propose a pattern-mixture type estimator for the CIF to evaluate the effect of a two-stage treatment regime. We derive the explicit form of its variance estimator and investigate the practical performance of this pattern-mixture

estimator using simulation. The first set of simulations are similar to those in Section 2.5, while the second set of simulations represent a case where the allocation probabilities may depend on some pre-determined covariate. We compare the performance of the proposed method to those of the inverse weighted and naïve approaches.

3.2 A PATTERN-MIXTURE TYPE ESTIMATOR

Here we assume the same set-up as described in Section 2.2, where we focus on the four treatment regimes that start with the initial treatment A_1, namely $A_1B_kB_l^{'}$, $k, l = 1, 2$. The notation remains the same as before. If we think about the cause 1 CIF for the regime $A_1B_kB_l^{'}$ as the conditional probability of a patient experiencing a type-1 event given the patient following the specific regime, then it can be decomposed into the following two parts based on the fact that the patient must have received either A_1B_k, if the initial treatment was working, or $A_1B_l^{'}$, otherwise.

$$\begin{aligned} F_{1,A_1B_kB_l^{'}}(t) &= P(T \leq t, \epsilon = 1 | A_1B_kB_l^{'}) \\ &= P(R = 1)P(T \leq t, \epsilon = 1 | A_1B_k) + \{1 - P(R = 1)\}P(T \leq t, \epsilon = 1 | A_1B_l^{'}) \\ &= P(R = 1)F_{1,A_1B_k}(t) + \{1 - P(R = 1)\}F_{1,A_1B_l^{'}}(t), \end{aligned} \tag{3.2.1}$$

where F_{1,A_1B_k} and $F_{1,A_1B_l^{'}}$ are the CIFs for treatment sequences of A_1B_k and $A_1B_l^{'}$, which can be estimated directly based on the subjects who received these treatment combinations. That is, to estimate F_{1,A_1B_k}, the standard non-parametric estimator of the CIF can be constructed using only the subjects who received A_1, responded to it and then received B_k. Suppose there are M distinct event times for these subjects, $t_1 < t_2 < \cdots < t_M$. Let Y_i be the number of subjects at risk, d_i be the number of subjects with the occurrence of the event of interest, and r_i be the number of subjects having the competing event at time t_i,

$i = 1, \ldots, M$. Then the standard non-parametric estimator of the CIF is:

$$\begin{aligned}\hat{F}_{1,A_1B_k}(t) &= \sum_{t_i \le t} \frac{d_i}{Y_i} \left\{ \prod_{j=1}^{i-1} \left(1 - \frac{d_j + r_j}{Y_j} \right) \right\} \\ &= \sum_{t_i \le t} \hat{S}_{A_1B_k}(t_i-) \frac{d_i}{Y_i},\end{aligned}$$

for $t_1 \le t$ and 0 otherwise, where $\hat{S}$ is the standard Kaplan-Meier estimator of the survival function for time to first event, either cause 1 or cause 2. $\hat{F}_{1,A_1B_l'}(t)$ can be constructed similarly.

For simplicity, we denote $P(R = 1) = pr$. This probability can easily be estimated by the sample proportion, $\hat{pr}$, which is the ratio of the number of responders to the number of subjects who were initially assigned to A_1. Plugging $\hat{pr}$ into the decomposition in (3.2.1) gives rise to the following estimator of the CIF for the treatment regime $A_1B_kB_l'$:

$$\hat{F}^{pm}_{1,A_1B_kB_l'}(t) = \hat{pr}\hat{F}_{1,A_1B_k}(t) + (1 - \hat{pr})\hat{F}_{1,A_1B_l'}(t), \tag{3.2.2}$$

which is referred to as the pattern-mixture estimator of CIF (PMCIF) in the sequel. Estimating the variance of $\hat{F}^{pm}_{1,A_1B_kB_l'}(t)$ is straightforward because the standard non-parametric estimators for $\hat{F}_{1,A_1B_k}(t)$ and $\hat{F}_{1,A_1B_l'}(t)$ have explicit variance estimators and $\hat{F}^{pm}_{1,A_1B_kB_l'}(t)$ is simply a linear combination of the two independent estimators. The estimated variance of $\hat{F}_{1,A_1B_k}(t)$ is given in Klein and Moeschberger (2003) as

$$\begin{aligned}\hat{\sigma}^2 \left(\hat{F}_{1,A_1B_k}(t) \right) &= \sum_{t_i \le t} \Bigg(\hat{S}_{A_1B_k}(t_i)^2 \left\{ \left(\hat{F}_{1,A_1B_k}(t) - \hat{F}_{1,A_1B_k}(t_i) \right)^2 \frac{r_i + d_i}{Y_i^2} \right\} \\ &\quad + \hat{S}_{A_1B_k}(t_i)^2 \left[\left\{ 1 - 2 \left(\hat{F}_{1,A_1B_k}(t) - \hat{F}_{1,A_1B_k}(t_i) \right) \right\} \frac{d_i}{Y_i^2} \right] \Bigg).\end{aligned}$$

The variance estimator of $\hat{F}_{1,A_1B_l'}(t)$ follows similarly. Therefore, the variance of $\hat{F}^{pm}_{1,A_1B_kB_l'}$ can be estimated as follows:

$$\hat{\sigma}^2 \left(\hat{F}^{pm}_{1,A_1B_kB_l'}(t) \right) = \hat{pr}^2 \hat{\sigma}^2 \left(\hat{F}_{1,A_1B_k}(t) \right) + (1 - \hat{pr})^2 \hat{\sigma}^2 \left(\hat{F}_{1,A_1B_l'}(t) \right). \tag{3.2.3}$$

It is worthwhile to note that the proposed pattern-mixture approach and inverse weighting approaches are equivalent for a continuous outcome Y without censoring. To demonstrate

this, assume we are interested in the mean of this outcome for a specific dynamic treatment regime $A_1B_kB_l^{'}$. Then, the pattern-mixture approach would suggest we estimate this mean outcome as:

$$\begin{aligned}
\bar{Y}^{pm} &= \hat{pr}\frac{1}{n_{A_1B_k}}\sum_{i\epsilon A_1B_k} Y_i + (1-\hat{pr})\frac{1}{n_{A_1B_l^{'}}}\sum_{i\epsilon A_1B_l^{'}} Y_i \\
&= \hat{pr}\frac{1}{n\hat{pr}\hat{\pi}_{B_k}}\sum_{i\epsilon A_1B_k} Y_i + (1-\hat{pr})\frac{1}{n(1-\hat{pr})\hat{\pi}_{B_l^{'}}}\sum_{i\epsilon A_1B_l^{'}} Y_i \\
&= \frac{1}{n}\left\{\sum_{i\epsilon A_1B_k}\frac{Y_i}{\hat{\pi}_{B_k}} + \sum_{i\epsilon A_1B_l^{'}}\frac{Y_i}{\hat{\pi}_{B_l^{'}}}\right\} = \bar{Y}^{ew},
\end{aligned}$$

where n, $n_{A_1B_k}$ and $n_{A_1B_l^{'}}$ are the number of subjects receiving A_1, A_1B_k and $A_1B_l^{'}$ and $\bar{Y}^{pm}$ and $\bar{Y}^{ew}$ represent the mean outcomes calculated using the pattern mixture approach and inverse weighting respectively. However, the equivalence of the two approaches can not be established analytically for time-to-event data where the event time is exposed to competing risks. In the following section, we will show that the two approaches generate almost equivalent results under various settings.

3.3 SIMULATION

3.3.1 A Simple Two Stage Randomization

Simulation studies were carried out to compare the proposed estimator to the naïve estimator and the inverse weighted estimators under different conditions. Again, only the subjects who received the initial treatment A_1 were considered since the data obtained from these subjects are independent of the data obtained from the subjects that were assigned to the other initial treatment A_2. The comparisons were made under scenarios 2 and 3 as in Section 2.5, namely, the data were generated at which 2) $Z_{1i} \sim 2 - Bernoulli(0.3)$ and $Z_{2i} \sim 2 - Bernoulli(0.3)$, and 3) $Z_{1i} = 1$ and $Z_{2i} \sim 2 - Bernoulli(0.5)$. Every model was repeated for $n = 300$ and for cases $R_i \sim Bernoulli(0.4)$ and $R_i \sim Bernoulli(0.7)$ and 4000 data sets were generated for each setting.

For each combination $\{(T_i^R, R_i, Z_{1i}, Z_{2i}, V_i, \Delta_i), i = 1, \ldots, n\}$ were generated using the same set-up as in Section 2.5 and only the results for the regimes $A_1B_1B_1^{'}$ and $A_1B_1B_2^{'}$ are given since the results for other regimes were similar.

In the following tables CI(t) stands for the naïve estimate of the CIF evaluated at time t, WCI(t) stands for the inverse probability weighted estimate of the CIF with fixed weights and WCI2(t) is the inverse weighted CIF estimate with estimated fixed weights , TWCI(t) is the estimate of the CIF with time-dependent weights, TWCI2(t) is the CIF estimate with estimated time-dependent weights, and PMCIF(t) is the pattern-mixture CIF estimate given in (3.2.2).

Table 4 shows the simulation results for Scenario-2 and Scenario-3 with sample size $n = 300$. For Scenario-2 all the estimators perform similarly for the first regime $A_1B_1B_1^{'}$ where they all produce unbiased estimates for the CIF at given times and they all have coverage rates close to the desired value of 95%. For the PMCIF estimator it can be seen that the straightforward variance estimator is working perfectly for both response rates. The similarity of all estimators including the naïve one is expected under this scenario due to the randomization probabilities being equal, i.e. $\pi_{B_1} = \pi_{B_1^{'}} = 0.3$. For the second regime $A_1B_1B_2^{'}$, the naïve estimator produces highly biased estimates which result in poor coverage rates where all the other estimators are still unbiased and have desirable coverage rates for both time points and both response rates.

For Scenario-3, the difference of the naïve estimator versus the inverse weighted and pattern-mixture estimators can be seen for both regimes at both time points. The naïve estimator stays biased with poor coverage rates under all conditions, while the pattern-mixture estimator is unbiased with coverage rates close to 95% just like the inverse probability weighted estimators.

3.3.2 Two Stage Randomization with Covariate Dependent Allocation

In practice, there may be information available suggesting that some treatment options are more effective in treating certain diseases for certain subgroups of patients. If such information is available, researchers may be inclined to design their experiments taking this

Table 4: Pattern-mixture- Scenario-2 and Scenario-3 with n=300: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation (σ), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

				$A_1B_1B_1'$					$A_1B_1B_2'$				
Scenario	t	pr	Method	True	Mean	σ	$\hat{\sigma}$	Cov	True	Mean	σ	$\hat{\sigma}$	Cov
2	0.5	0.4	CI	0.185	0.185	0.041	0.041	0.939	0.248	0.283	0.036	0.036	0.863
			WCI		0.185	0.041	0.040	0.934		0.248	0.035	0.034	0.937
			WCI2		0.185	0.041	0.040	0.934		0.248	0.035	0.034	0.940
			TWCI		0.185	0.041	0.040	0.940		0.248	0.035	0.034	0.942
			TWCI2		0.185	0.041	0.040	0.938		0.248	0.035	0.034	0.942
			PMCIF		0.185	0.041	0.040	0.937		0.248	0.035	0.034	0.940
		0.7	CI	0.158	0.158	0.038	0.038	0.943	0.190	0.229	0.038	0.038	0.844
			WCI		0.158	0.038	0.038	0.938		0.190	0.036	0.035	0.935
			WCI2		0.158	0.038	0.038	0.938		0.190	0.036	0.035	0.938
			TWCI		0.158	0.038	0.037	0.940		0.190	0.035	0.034	0.936
			TWCI2		0.158	0.038	0.038	0.940		0.190	0.035	0.035	0.937
			PMCIF		0.158	0.038	0.038	0.939		0.190	0.035	0.035	0.937
	1	0.4	CI	0.342	0.343	0.051	0.052	0.950	0.389	0.425	0.039	0.042	0.891
			WCI		0.343	0.051	0.049	0.938		0.390	0.041	0.040	0.944
			WCI2		0.343	0.051	0.049	0.941		0.390	0.041	0.040	0.946
			TWCI		0.343	0.051	0.050	0.942		0.390	0.041	0.041	0.954
			TWCI2		0.343	0.051	0.050	0.946		0.390	0.040	0.041	0.954
			PMCIF		0.343	0.051	0.051	0.950		0.390	0.040	0.042	0.957
		0.7	CI	0.306	0.307	0.049	0.050	0.947	0.330	0.370	0.043	0.045	0.882
			WCI		0.307	0.049	0.048	0.934		0.331	0.045	0.043	0.937
			WCI2		0.307	0.049	0.048	0.937		0.331	0.044	0.043	0.939
			TWCI		0.307	0.049	0.048	0.940		0.331	0.044	0.044	0.942
			TWCI2		0.307	0.049	0.048	0.942		0.330	0.044	0.044	0.943
			PMCIF		0.307	0.049	0.049	0.944		0.331	0.044	0.045	0.948
3	0.5	0.4	CI	0.185	0.170	0.026	0.026	0.883	0.248	0.215	0.028	0.029	0.779
			WCI		0.185	0.029	0.029	0.939		0.248	0.032	0.032	0.946
			WCI2		0.185	0.029	0.029	0.940		0.249	0.032	0.032	0.949
			TWCI		0.185	0.029	0.029	0.939		0.248	0.032	0.032	0.942
			TWCI2		0.185	0.029	0.029	0.940		0.248	0.032	0.032	0.948
			PMCIF		0.185	0.030	0.029	0.942		0.248	0.032	0.033	0.952
		0.7	CI	0.157	0.147	0.022	0.022	0.908	0.190	0.166	0.023	0.023	0.816
			WCI		0.158	0.025	0.025	0.935		0.190	0.027	0.027	0.940
			WCI2		0.158	0.025	0.025	0.935		0.190	0.027	0.027	0.940
			TWCI		0.158	0.025	0.025	0.938		0.190	0.027	0.027	0.939
			TWCI2		0.158	0.025	0.025	0.939		0.190	0.027	0.027	0.941
			PMCIF		0.158	0.025	0.025	0.936		0.190	0.027	0.027	0.941
	1	0.4	CI	0.342	0.322	0.032	0.034	0.905	0.389	0.356	0.033	0.035	0.838
			WCI		0.342	0.035	0.035	0.946		0.389	0.036	0.036	0.946
			WCI2		0.342	0.035	0.035	0.946		0.389	0.036	0.036	0.948
			TWCI		0.342	0.035	0.035	0.946		0.389	0.036	0.036	0.946
			TWCI2		0.342	0.035	0.035	0.947		0.389	0.036	0.036	0.949
			PMCIF		0.342	0.035	0.037	0.958		0.389	0.036	0.038	0.960
		0.7	CI	0.306	0.292	0.028	0.029	0.925	0.330	0.306	0.028	0.030	0.880
			WCI		0.307	0.031	0.031	0.945		0.330	0.031	0.031	0.950
			WCI2		0.307	0.031	0.031	0.944		0.330	0.031	0.031	0.949
			TWCI		0.307	0.031	0.031	0.945		0.330	0.031	0.031	0.948
			TWCI2		0.307	0.031	0.031	0.946		0.330	0.031	0.031	0.950
			PMCIF		0.307	0.031	0.032	0.953		0.330	0.031	0.032	0.956

information into account and randomizing patients into treatments which they may be more likely to benefit from. To investigate the performance of the proposed estimator under such a setting, a covariate X was generated from a $Bernoulli(0.5)$ to represent the gender of a patient. As in previous section only the subjects who received the initial treatment A_1 were considered and all simulations were repeated 4000 times for cases $n = 300$ and $R_i \sim Bernoulli(0.4)$ and $R_i \sim Bernoulli(0.7)$. Comparisons then were made under two different scenarios.

For scenario-A, treatments B_1 and $B_1^{'}$ are assumed to have higher chances of success for females. For this reason, the females were given higher probabilities of being assigned to these treatments depending on their response status. For men, the allocation was carried out randomly with equal probability since no such information was available. If we let Z_{1f} and Z_{2f} be the assignment indicators for females and Z_{1m} and Z_{2m} assignment indicators for males, the randomization was carried out using $Z_{1f} \sim 2 - Bernoulli(0.75)$, $Z_{2f} \sim 2 - Bernoulli(0.75)$ and $Z_{1m} \sim 2 - Bernoulli(0.5)$, $Z_{2m} \sim 2 - Bernoulli(0.5)$.

The times to response, T_i^R, were generated from Exponential (0.20) and restricted at 1 year. The times to death from the second randomization ($T^*_{A_1B_{ki}}$ or $T^*_{A_1B^{'}_{li}}$, $k, l = 1, 2$) were drawn from exponential distributions with different parameter values for females and males. For females, the times to death were generated from exponential distributions with parameter values of 1.25 for the sequence of treatments A_1B_1, 0.75 for the A_1B_2, 0.65 for the $A_1B_1^{'}$ and 0.25 for the $A_1B_2^{'}$ treatments. For males, the times to death was generated from exponential distributions with parameter values of 0.90 for the sequence of treatments A_1B_1, 0.70 for the A_1B_2, 0.50 for the $A_1B_1^{'}$ and 0.40 for the $A_1B_2^{'}$ treatments. Censoring time was generated from $Uniform(1.5, 2)$ which resulted in approximately 12% censoring for $P(R = 1) = 0.4$ and 16% censoring for $P(R = 1) = 0.7$ and the overall survival time was calculated as given in Section 2.5 for each subject i.

For scenario-B, it is assumed that B_1 works better for female responders than B_2, and $B_2^{'}$ works better for female non-responders than $B_1^{'}$, while the only information on males is that $B_2^{'}$ works better for male non-responders than $B_1^{'}$. Therefore females were assigned to B_1 and $B_2{'}$ and males were assigned to $B_2{'}$ with higher probabilities. More specifically

$Z_{1f} \sim 2 - Bernoulli(0.75)$, $Z_{2f} \sim 2 - Bernoulli(0.30)$ and $Z_{1m} \sim 2 - Bernoulli(0.50)$, $Z_{2m} \sim 2 - Bernoulli(0.30)$.

The times to response, T_i^R, were generated from Exponential (0.20) and restricted at 1 years. The times to death from the second randomization were again drawn from different exponential distributions with different the parameter values for females and males. For females, the times to death were generated from exponential distributions with parameter values of 1.25 for the sequence of treatments A_1B_1, 0.75 for the A_1B_2, 0.25 for the A_1B_1' and 0.65 for the A_1B_2' treatments. For males, the times to death was generated from exponential distributions with parameter values of 0.90 for the sequence of treatments A_1B_1, 0.70 for the A_1B_2, 0.30 for the A_1B_1' and 0.70 for the A_1B_2' treatments. Censoring time was generated from $Uniform(1.5, 2)$ which resulted in approximately 13% censoring for $P(R = 1) = 0.4$ and 17% censoring for $P(R = 1) = 0.7$ and the overall survival time was calculated as given in Section 2.5 for each subject i.

Table 5 shows the simulation results for Scenario-A and Scenario-B with sample size $n = 300$. For Scenario-A, the naïve estimator produces biased estimates for both regimes under both response rates, performing worse for the second regime. The inverse probability weighted estimators and the pattern-mixture estimator all perform as desired under all cases and for both time points. For Scenario-B even though the naïve estimator is unbiased for the second regime, it still produces biased estimates for the first one. The proposed estimators are again all unbiased and have much better coverage rates. In general, it can be seen that under some cases, the variance estimator of the pattern-mixture estimator may tend to slightly under-estimate the true variance, resulting in slightly lower coverage rates compared to the inverse-weighted methods but overall the proposed methods are all highly preferable to the naïve approach.

Table 5: Pattern-mixture- Scenario-A and Scenario-B with n=300, covariate dependent allocation: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation (σ), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

				$A_1B_1B_1'$					$A_1B_1B_2'$				
Scenario	*t*	*pr*	Method	True	Mean	σ	$\hat{\sigma}$	Cov	True	Mean	σ	$\hat{\sigma}$	Cov
A	0.5	0.4	CI	0.172	0.168	0.028	0.027	0.934	0.224	0.196	0.033	0.033	0.837
			WCI		0.171	0.029	0.028	0.937		0.224	0.039	0.039	0.942
			WCI2		0.171	0.029	0.028	0.938		0.224	0.039	0.039	0.942
			TWCI		0.171	0.029	0.028	0.940		0.223	0.039	0.038	0.940
			TWCI2		0.171	0.029	0.028	0.940		0.223	0.039	0.039	0.941
			PMCIF		0.171	0.029	0.028	0.938		0.224	0.039	0.039	0.940
		0.7	CI	0.149	0.145	0.025	0.026	0.938	0.175	0.155	0.028	0.028	0.860
			WCI		0.149	0.027	0.026	0.940		0.175	0.034	0.033	0.935
			WCI2		0.149	0.027	0.026	0.940		0.175	0.034	0.033	0.936
			TWCI		0.148	0.026	0.026	0.940		0.175	0.034	0.033	0.934
			TWCI2		0.148	0.026	0.026	0.940		0.175	0.033	0.033	0.935
			PMCIF		0.149	0.027	0.026	0.940		0.175	0.034	0.031	0.922
	1	0.4	CI	0.325	0.321	0.035	0.036	0.954	0.371	0.341	0.041	0.042	0.875
			WCI		0.326	0.036	0.035	0.945		0.370	0.045	0.043	0.936
			WCI2		0.326	0.036	0.035	0.946		0.370	0.045	0.044	0.937
			TWCI		0.326	0.035	0.035	0.948		0.370	0.045	0.044	0.941
			TWCI2		0.326	0.035	0.035	0.950		0.370	0.045	0.044	0.940
			PMCIF		0.326	0.036	0.036	0.951		0.371	0.045	0.044	0.943
		0.7	CI	0.293	0.288	0.033	0.034	0.952	0.316	0.293	0.035	0.037	0.896
			WCI		0.293	0.034	0.034	0.947		0.316	0.039	0.039	0.944
			WCI2		0.293	0.034	0.034	0.947		0.316	0.039	0.039	0.943
			TWCI		0.293	0.034	0.034	0.946		0.316	0.040	0.039	0.942
			TWCI2		0.293	0.034	0.034	0.948		0.316	0.039	0.039	0.944
			PMCIF		0.293	0.034	0.034	0.949		0.316	0.039	0.038	0.932
B	0.5	0.4	CI	0.238	0.202	0.035	0.034	0.806	0.158	0.158	0.026	0.026	0.946
			WCI		0.238	0.041	0.041	0.941		0.158	0.026	0.026	0.945
			WCI2		0.238	0.041	0.041	0.939		0.158	0.026	0.026	0.944
			TWCI		0.237	0.041	0.041	0.939		0.158	0.026	0.026	0.945
			TWCI2		0.237	0.041	0.041	0.940		0.158	0.026	0.026	0.945
			PMCIF		0.238	0.042	0.041	0.937		0.158	0.026	0.026	0.942
		0.7	CI	0.182	0.155	0.028	0.029	0.816	0.142	0.141	0.025	0.025	0.939
			WCI		0.182	0.034	0.034	0.946		0.142	0.025	0.025	0.937
			WCI2		0.182	0.034	0.034	0.944		0.142	0.025	0.025	0.939
			TWCI		0.182	0.034	0.034	0.942		0.142	0.025	0.025	0.938
			TWCI2		0.182	0.034	0.034	0.944		0.142	0.025	0.025	0.942
			PMCIF		0.182	0.034	0.033	0.931		0.142	0.025	0.025	0.938
	1	0.4	CI	0.381	0.341	0.042	0.044	0.844	0.309	0.308	0.033	0.034	0.953
			WCI		0.380	0.047	0.045	0.940		0.308	0.033	0.033	0.944
			WCI2		0.382	0.047	0.046	0.939		0.308	0.033	0.033	0.943
			TWCI		0.379	0.047	0.046	0.942		0.308	0.033	0.033	0.946
			TWCI2		0.379	0.047	0.046	0.944		0.308	0.033	0.033	0.945
			PMCIF		0.380	0.047	0.046	0.941		0.308	0.033	0.033	0.949
		0.7	CI	0.321	0.291	0.035	0.037	0.866	0.285	0.283	0.031	0.033	0.957
			WCI		0.321	0.040	0.040	0.948		0.285	0.032	0.033	0.952
			WCI2		0.321	0.039	0.040	0.950		0.285	0.032	0.033	0.953
			TWCI		0.321	0.040	0.040	0.949		0.285	0.032	0.033	0.952
			TWCI2		0.321	0.040	0.040	0.952		0.285	0.032	0.033	0.953
			PMCIF		0.321	0.040	0.039	0.941		0.285	0.032	0.033	0.954

3.4 DATA ANALYSIS

The proposed pattern-mixture type estimator was applied to the data set from the Cancer and Leukemia Group B (CALGB) trial (Stone et al., 1995) which was explained in Section 2.7. The pseudo event types created in Section 2.7 were recorded and for the purpose of demonstrating this method and comparing it to the inverse weighted and the naïve methods, the same pseudo event type values were used. We calculated the standard CIF estimator (CI), the CIF estimator with fixed weights (WCI) and the pattern-mixture type CIF estimator (PMCIF) for the cause 1 event of the four regimes. The results are shown in Figure 4.

It can be seen from Figure 4 that the standard CIF estimator generally produces higher estimates for the CIF and the inverse weighted estimator and the pattern-mixture estimator produce very similar results. We can also obtain a weighted estimate for the overall survival function based on the estimated cause-1 and cause-2 PMCIFs. That is,

$$\hat{S}^{pm}_{A_kB_l}(t) = 1 - \hat{F}^{pm}_{1,A_kB_l}(t) - \hat{F}^{pm}_{2,A_kB_l}(t),$$

for $k, l = 1, 2$. The survival probabilities for the four regimes calculated using the above formula are plotted in Figure 5. It can be seen that the estimated survival probabilities using the pattern-mixture approach are quite similar to the survival probabilities obtained using the inverse weighting approach which are plotted in Figure 3. As before, none of the treatment regimes clearly out perform the others.

3.5 SUMMARY

In this chapter, we proposed a pattern-mixture type estimator for the CIF which can be used when a two-stage randomization is carried out and the outcome of interest is subject to competing-risk censoring. The application of the proposed method is straightforward using existing packages in R and it has an explicit variance estimator making inference simple. We showed that the proposed estimator produces unbiased estimates for the CIF

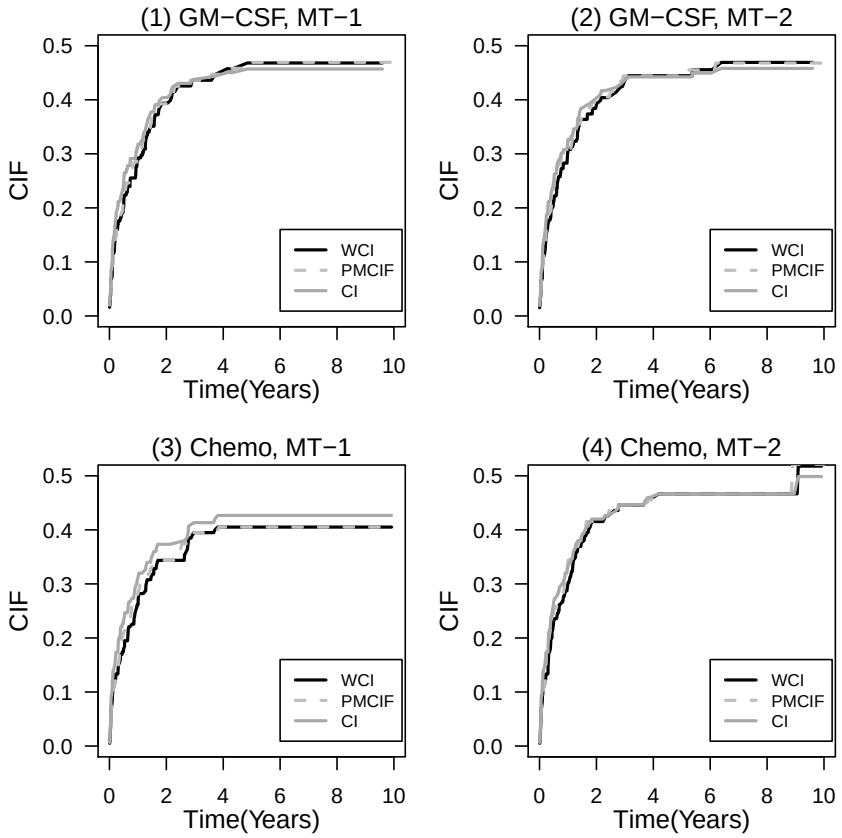

Figure 4: Pattern-mixture, inverse weighted and naïve CIF estimators for the four regimes.

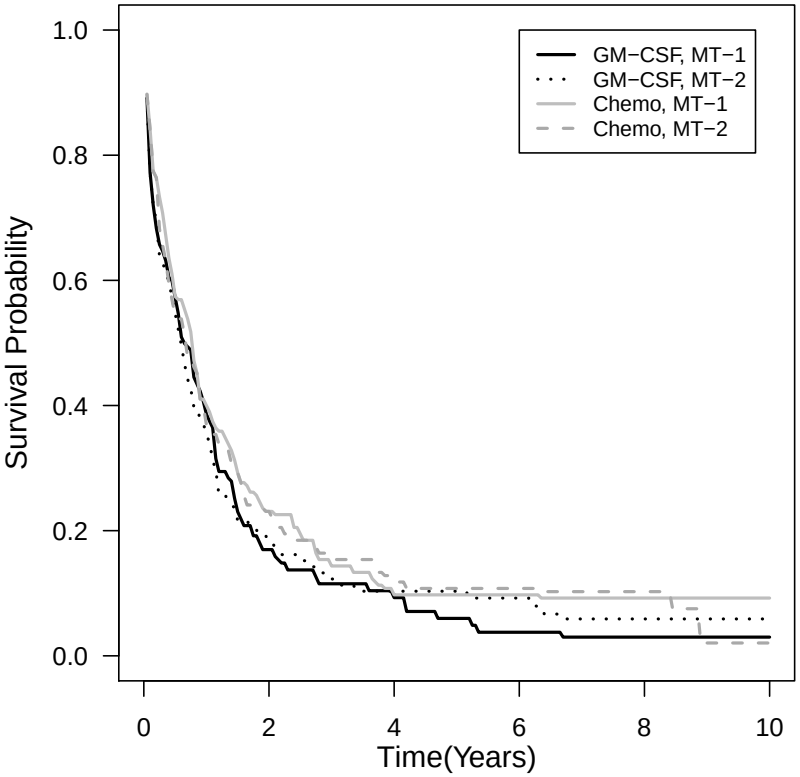

Figure 5: Comparison of regimes with respect to the pattern-mixture survival estimates.

with good coverage rates where the naïve estimator fails to do so. In general, it can be seen that the pattern-mixture estimator produces very similar results to those obtained by inverse weighting even when a covariate is involved in the allocation process. One possible limitation of this method is that the sample size needs to be sufficiently large within each stratum, as in any stratified data analysis technique. When there are only a few subjects in a stratified group, the estimation of the CIF may not be possible or might produce an unreliable result. However, if the sample size within each stratum is sufficient, it is clearly seen that the pattern-mixture approach provides a good alternative to the inversely weighted methods and it will be useful in estimation of the CIF when dynamic treatment regimes are involved.

4.0 REGRESSION ANALYSIS FOR THE CUMULATIVE INCIDENCE FUNCTION UNDER TWO-STAGE RANDOMIZATION

4.1 INTRODUCTION

When practitioners deal with time-to-event data, it is always of interest to model or control for the covariate effects when they are present. Methods that are used for such modeling purposes have been greatly studied by researchers. The Cox regression model (Cox, 1972) and the accelerated failure model (Wei et al., 1990; Jin et al., 2003, among others) are two popular methods of modeling covariate effects on survival outcomes.

Over the years, competing risks became more recognized by practitioners and new methods were developed to deal with such time-to-event data where the event time is dependently censored by some competing events. Fine and Gray (1999) extended the Cox regression model to handle competing risks. They proposed a proportional hazard model for hazard of sub-distribution which is more commonly called "cumulative incidence function (CIF)." Fine (2001) proposed a semi-parametric regression for CIF extending the standard log-linear regression model to competing risks settings. Klein and Andersen (2005) proposed a different model which is based on pseudo values of CIF. Jeong and Fine (2007) introduced a parametric regression model for CIF using the two-parameter Gompertz model as baseline CIF. Shi et al. (2013) later developed a modified three-parameter logistic regression model that outperforms the Gompertz model especially when CIF curves have sigmoidal shapes. Scheike et al. (2008) extended the Fine and Gray model using direct binomial regression so that it can handle time-varying coefficients. There is also extensive work on regression analysis of competing risks data based on the cause-specific hazard functions (Korn and Dorey,

1992; Cheng et al., 1998; Andersen et al., 2002; Hyun et al., 2009). In this chapter, we focus on modeling the CIF, since it has a nice probability interpretation and is more appealing to practitioners.

Although the regression models for time-to-event data have been studied in detail for both regular and competing-risk censoring, there has not been much research for the two-stage randomization setting. Murphy (2003) used experimental or observational data to estimate the best regime that results in a maximal mean response. Zhao et al. (2009) developed a learning method utilizing Q-learning to choose an optimal policy from a single training data set. Henderson et al. (2010) proposed a modeling and estimation strategy that incorporates the regret functions of Murphy (2003) into a regression model for observed responses. More recently, Goldberg and Kosorok (2012) proposed a method that used survival as an outcome and focused on finding the treatment policy that maximized the survival function. These methods focus on choosing the best available treatment regime that will give rise to the best outcome for an individual with certain characteristics, rather than modeling the effects of covariates on the survival function or in our case the CIF for a specific treatment regime. The latter is also of great practical interest, since when a patient is recommended to follow a specific regime, she may wonder how likely she would have the desirable outcome (e.g., no recurrence of breast cancer in five years), given her clinical characteristics. The question can be addressed by performing regression analysis using the data from the current two-stage randomization trial. The former endeavors of searching the best treatment regime are usually based on historical data. Another line of research on regression analysis for dynamic treatment regimes is to extend the Cox model (Cox, 1972) to two-stage randomization designs for comparisons of multiple treatment regimes (Lokhnygina and Helterbrand, 2007) which takes different treatment regimes as a single categorical covariate.

We focus on the approach of directly modeling covariate effects on a survival outcome. To our best knowledge, there has not been any publications modeling covariate effects on the CIFs of specific regimes in a two-stage randomized trial. Hence, in this chapter we will carry out various regression analyses to model covariate effects on the CIF of a specific treatment regime. More specifically, we will extend the multi-state (Cheng et al., 1998; Andersen et al.,

2002), Fine and Gray (1999) and Scheike et al. (2008) models for dynamic treatment regimes using inverse weighting and the pattern-mixture approach and provide detailed simulation results.

Throughout the chapter we assume there are only two causes of failure, cause 1 event is the event of interest, and a two-stage randomized trial is being carried out as in Section 2.2. Also, we assume only patients who are assigned to treatment A_1 in the initial assignment are being considered. In addition to the notation given in Section 2.2, in the following sections, let $\mathbf{X}$ be a covariate vector associated with the occurrence of type 1 event. We define the conditional CIF given $\mathbf{X}$ for treatment regime $A_1B_kB_l^{'}$ as $F_{1,A_1B_kB_l^{'}}(t;\mathbf{X}) = P(T \leq t, \Delta = 1|\mathbf{X}, \text{regime} = A_1B_kB_l^{'})$.

4.2 THE MULTI-STATE MODEL

4.2.1 The Existing Model

When competing risks are present, the cause-specific CIF can be estimated using the multi-state model (Cheng et al., 1998; Andersen et al., 2002). More specifically, the Cox model is applied to evaluate the covariate effects on the cause-specific hazard function. The cause-k ($k = 1, 2$) hazard function can be written as

$$\lambda_k(t) = \lim_{h \to 0} \frac{1}{h} P(t \leq T < t + h, \Delta = k | T \geq t).$$

Under the multi-state model it is assumed

$$\lambda_k(t; \mathbf{X}) = \lambda_{0k}(t) \exp(\beta_k^{'} \mathbf{X}), \tag{4.2.1}$$

where β_k is a vector of unknown regression parameters and $\lambda_{0k}(.)$ is a nonnegative but otherwise unspecified baseline function for the cause k event. The regression coefficients β_k can be estimated using a partial likelihood principle by treating all the non cause-k failures

as censored observations. Let $N_{ki}(t) = I(V_i \leq t, \Delta_i = k)$ and $Y_i(t) = I(V_i \geq t)$. Then the partial likelihood for β_k can be written as

$$L(\beta_k) = \prod_t \prod_{i=1}^{n} \left(\frac{\exp(\beta_k' \mathbf{X}_i)}{\sum_{j=1}^{n} \exp(\beta_k' \mathbf{X}_j) Y_j(t)} \right)^{\Delta N_{ki}(t)}, \tag{4.2.2}$$

where the first product is over all distinct event times and the second product is over all subjects. Maximizing this likelihood, it is straightforward to obtain estimates for the regression coefficients. If we let $\hat{\beta}_k$ be the maximum partial likelihood estimates and $\hat{\Lambda}_k(t)$ be the Breslow (Breslow, 1974) estimate for the cumulative cause-k hazard function, then the cause-1 CIF can be estimated by:

$$\hat{F}_1(t; \mathbf{X}) = \int_0^t \hat{S}(u; \mathbf{X}) d\hat{\Lambda}_1(u; \mathbf{X}) = \sum_{i=1}^{n} \frac{\hat{S}(V_i; \mathbf{X}) I(V_i \leq t, \Delta_i = 1) \exp(\hat{\beta}_1' \mathbf{X})}{\sum_{j=1}^{n} I(V_i \leq V_j) \exp(\hat{\beta}_1' \mathbf{X_j})}, \tag{4.2.3}$$

where $\hat{S}(u; \mathbf{X}) = \exp\{-\sum_{k=1}^{2} \hat{\Lambda}_k(u; \mathbf{X})\}$, $\hat{\Lambda}_k(u; x) = \hat{\Lambda}_{0k}(u) \exp(\hat{\beta}_k' \mathbf{X})$ and

$$\hat{\Lambda}_{0k}(t) = \frac{\sum_{i=1}^{n} I(V_i \leq t, \Delta_i = k)}{\sum_{j=1}^{n} I(V_i \leq V_j) \exp(\hat{\beta}_k' \mathbf{X_j})}.$$

4.2.2 Extension Using Inverse Weighting

The above likelihood is constructed under the assumption that all the subjects who failed from cause-1 should contribute to the estimation process equally, independent of the treatment sequence they receive. Under a two-stage randomization setting, when estimating the regression coefficients for a specific regime, subjects should contribute to this likelihood proportional to their inverse probability weights as in Chapter 2. For this reason, for regime $A_1 B_k B_l'$ we propose maximizing the below likelihood for estimating the regression coefficients associated with cause-1:

$$L_{A_1 B_k B_l'}(\beta_1) = \prod_t \prod_{i=1}^{n} \left(\frac{\exp(\beta_1' \mathbf{X}_i)}{\sum_{j=1}^{n} \exp(\beta_1' \mathbf{X}_j) Y_j^w(t)} \right)^{\Delta N_{1i}^w(t)}, \tag{4.2.4}$$

where $N_{1i}^w(t) = I(V_i \leq t, \Delta_i = 1) \hat{Q}_{A_1 B_k B_l', i}$, $Y_i^w(t) = I(V_i \geq t) \hat{Q}_{A_1 B_k B_l', i}$ and $\hat{Q}_{A_1 B_k B_l', i} = R_i I(Z_{1i} = k)/\hat{\pi}_{B_k} + (1 - R_i) I(Z_{2i} = l)/\hat{\pi}_{B_l'}$ as in Chapter 2. Denote the regression coefficients

obtained by maximizing this weighted likelihood by $\hat{\beta}_1^w$. Then, the CIF for regime $A_1B_kB_l^{'}$ can be estimated by

$$\hat{F}_1^w(t;\mathbf{X}) = \int_0^t \hat{S}^w(u;\mathbf{X}) d\hat{\Lambda}_1^w(u;\mathbf{X}), \tag{4.2.5}$$

where $\hat{S}^w(u;\mathbf{X}) = \exp\{-\sum_{k=1}^2 \hat{\Lambda}_k^w(u;\mathbf{X})\}$, $\hat{\Lambda}_k^w(u;\mathbf{X}) = \hat{\Lambda}_{0k}^w(u)\exp(\hat{\beta}_k^{w'}\mathbf{X})$ and

$$\hat{\Lambda}_{0k}^w(t) = \frac{\sum_{i=1}^n I(V_i \le t, \Delta_i = k)}{\sum_{j=1}^n I(V_i \le V_j)\exp(\hat{\beta}_k^{w'}\mathbf{X}_j)}.$$

The implementation of our approach is straightforward using existing R functions. The estimation of the regression coefficients is carried out by the function "coxph" in the **survival** package and the estimated CIF values are obtained with the help of the function "msfit" in the **mstate** package. Note that the difference between (4.2.2) and (4.2.4) is the additional layer of weighting $\hat{Q}_{A_1B_kB_l^{'},i}$. If we use the fixed weights as discussed in Sections 2.3.1 and 2.3.2, those who responded to the initial treatment A_1 and were assigned to the subsequent treatment B_k would receive the same weight, for example 3.3, while those who did not respond to A_1 and were assigned to $B_l^{'}$ would receive another weight, say 2. Based on these weights, we create an augmented data set by repeating each of the subjects consistent with A_1B_k 33 times and each of the subjects receiving $A_1B_l^{'}$ 20 times. The R function "coxph" is applied to analyze the augmented data, and the resulting standard deviations are adjusted by the squared root of the ratio of the number of subjects in the augmented data to the number of subjects who were initially assigned to A_1. However, for time-varying weights, the trick may not work. One can write down the likelihood function similar as (4.2.4) with time-varying weights, and obtain the maximum likelihood estimators using some numerical methods such as "multiroot" in R.

4.2.3 Pattern-mixture Extension

It is also possible to extend the pattern-mixture approach discussed in Chapter 3 to regression settings. Recall that the CIF of the regime $A_1B_kB_l^{'}$ can be thought of as a linear combination

of two separate models for patients who followed A_1B_k and patients who followed $A_1B_l^{'}$. Similar to the decomposition in (3.2.1), we can write:

$$F_{1,A_1B_kB_l^{'}}(t;\mathbf{X}) = P(R=1)F_{1,A_1B_k}(t;\mathbf{X}) + \{1-P(R=1)\}F_{1,A_1B_l^{'}}(t;\mathbf{X}). \tag{4.2.6}$$

Let $\hat{\beta}_{1k}$ be the regression coefficient estimates for treatment combination A_1B_k obtained by maximizing the likelihood in (4.2.2) for only subjects who received A_1B_k and let $\hat{F}_{1,A_1B_k}(t;\mathbf{X})$ be the CIF estimate calculated using only the same subset of subjects by plugging $\hat{\beta}_{1k}$ in (4.2.3). Likewise let $\hat{\beta}_{1l'}$ be the regression coefficient estimates for treatment combination $A_1B_l^{'}$ obtained by maximizing the likelihood in (4.2.2) for only subjects who received $A_1B_l^{'}$ and let $\hat{F}_{1,A_1B_l^{'}}(t;\mathbf{X})$ be the CIF estimate calculated using only the same subset of subjects by plugging $\hat{\beta}_{1l'}$ in (4.2.3). Then we propose the regime-specific CIF can be estimated as a linear combination of these two estimates:

$$\hat{F}_{1,A_1B_kB_l^{'}}(t;\mathbf{X}) = \hat{pr}\hat{F}_{1,A_1B_k}(t;\mathbf{X}) + (1-\hat{pr})\hat{F}_{1,A_1B_l^{'}}(t;\mathbf{X}), \tag{4.2.7}$$

where $\hat{pr}$ is simply the estimated response probability. The analysis in R can be carried out simply by dividing the data into two parts, the patients who received A_1B_k and the patients who received $A_1B_l^{'}$, and using the functions "coxph" and "msfit." After this, the resulting CIF estimates can be combined as in (4.2.7) and the variance estimates can be combined as in (3.2.3) to obtain the final results for the regime $A_1B_kB_l^{'}$.

4.3 THE FINE AND GRAY MODEL

4.3.1 The Existing Model

Another popular regression model for a survival outcome subject to competing-risk censoring was proposed by Fine and Gray (1999) which assumed:

$$g\{F_1(t;\mathbf{X})\} = h_0(t) + \beta^{'}\mathbf{X}, \tag{4.3.1}$$

where g is some fixed transformation function, $h_0(t)$ is an invertible and monotonically increasing function, and β is a parameter vector.

To directly infer the effects of covariates on the CIF, Fine and Gray (1999) considered the transformation $g(u) = \log\{-\log(1-u)\}$ which has a proportional hazards interpretation for sub-distribution. Define

$$\begin{aligned}\lambda_1^*(t;\mathbf{X}) &= \lim_{h\to 0}\frac{1}{h}P\{t \le T \le t+h, \Delta = 1 | T \ge t \cup (T \le t \cap \Delta \ne 1), \mathbf{X}\} \\ &= \frac{dF_1(t;\mathbf{X})/dt}{1-F_1(t;\mathbf{X})} = -\frac{d\log\{1-F_1(t;\mathbf{X})\}}{dt},\end{aligned} \tag{4.3.2}$$

which is called the hazard of sub-distribution in Gray (1988). λ_1^* can be thought of as the hazard function of the improper variable $T^* = I(\Delta = 1)T + (1 - I(\Delta = 1))\infty$. The risk set associated with this hazard function contains both the subjects who have already failed from cause 2 event before time t and those who have never failed from any event by t.

When right censoring is present, Fine and Gray (1999) adopted inverse probability censoring weighting (IPCW) in order to obtain an unbiased estimating function. They assumed censoring time C was independent of T, ϵ and X, and $P(C \ge t) = G(t)$. For subject i, they defined the vital status as $r_i(t) = I\{C_i \ge (T_i \wedge t)\}$, where $T_i \wedge t = \min(T_i, t)$, and the weight $w_i(t) = \frac{r_i(t)\hat{G}(t)}{\hat{G}(V_i \wedge t)}$, where $\hat{G}$ is the Kaplan-Meier estimator for the censoring survival time. Letting $N_{1i}(t) = I(T_i \le t, \Delta_i = 1)$ and $Y_i^*(t) = 1 - N_{1i}(t-)$, they proposed the following weighted score function for the censored data:

$$U(\beta) = \sum_{i=1}^{n}\int_0^\infty \left[\mathbf{X}_i - \frac{\sum_j w_j(s)Y_j^*(s)\mathbf{X}_j \exp(\beta'\mathbf{X}_j)}{\sum_j w_j(s)Y_j^*(s)\exp(\beta'\mathbf{X}_j)}\right] w_i(s)dN_{1i}(s). \tag{4.3.3}$$

4.3.2 Extensions Using Inverse Weighting and the Pattern-mixture Approach

When fitting the above regression model under a two-stage randomization setting, we propose that the individuals should contribute to the partial likelihood function proportional to their inverse probability weights defined as in Chapter 2 of this dissertation, therefore another

layer of weighting is required. Below is the formal representation of the proposed score function for the regime $A_1B_kB_l^{'}$:

$$U_{A_1B_kB_l^{'}}(\beta) = \sum_{i=1}^{n} \int_0^{\infty} \left[\mathbf{X}_i - \frac{\sum_j w_j(s) Y_j^{*w}(s) \mathbf{X}_j \exp(\beta^{'} \mathbf{X}_j)}{\sum_j w_j(s) Y_j^{*w}(s) \exp(\beta^{'} \mathbf{X}_j)} \right] w_i(s) dN_{1i}^{w}(s), \tag{4.3.4}$$

where $N_{1i}^{w}(t) = I(T_i \leq t, \Delta_i = 1)\hat{Q}_{A_1B_kB_l^{'},i}$, $Y_i^{*w}(t) = \{1 - I(T_i \leq t, \Delta_i = 1)\}\hat{Q}_{A_1B_kB_l^{'},i}$ and $\hat{Q}_{A_1B_kB_l^{'},i}$ are as defined in Chapter 2.

The implementation of our approach is again straightforward using an existing R function, "crr" in the **cmprsk** package. Note that the difference between (4.3.3) and (4.3.4) is just the additional layer of weighting $\hat{Q}_{A_1B_kB_l^{'},i}$. Based on these weights, we again create an augmented data set by repeating each of the subjects consistent with A_1B_k and each of the subjects receiving $A_1B_l^{'}$ proportional to their weights. The R function "crr" is then applied to analyze the augmented data to obtain estimates for the regression coefficients and the CIF. The variance for the CIF estimate is achieved by bootstrapping.

We can also assume the Fine and Gray model for both treatment sequences A_1B_k and $A_1B_l^{'}$. Then, the CIF estimates for a specific regime can be obtained by splitting the data consistent with the regime into two parts depending on the treatment combinations they received and then fitting the Fine and Gray model to each subset of data. Then the estimates from these two models can be combined as in (4.2.7) to get the pattern-mixture CIF estimate for the regime. Variance estimates from bootstrapping can also be combined as in (3.2.3) to obtain a final variance estimate.

4.4 THE SCHEIKE ET AL. MODEL

4.4.1 The Existing Model

Our extensions to the Fine and Gray model are natural when we think about the data structure of a two-stage randomization setting. However, the proportionality assumption for sub-distribution hazards in the Fine and Gray model may be too restrictive for a two-stage randomized trial. During the course of follow-up, the treatment assignment is changed

depending on a subject's response to the initial treatment. As a result, the covariate effects on the CIF may also change. The following generalized model proposed in Scheike et al. (2008) may be more desirable for the two-stage randomization setting:

$$g\{F_1(t;\mathbf{X})\} = \mathbf{X}_1'\boldsymbol{\alpha}(t) + h(\mathbf{X}_2, \boldsymbol{\gamma}, t),$$

where g and h are known link functions, $\boldsymbol{\alpha}(t)$ are time-varying coefficients for a sub-vector of covariates $\mathbf{X}_1$, and $\mathbf{X}_2$ are the remaining covariates in $\mathbf{X}$ with time-independent coefficients $\boldsymbol{\gamma}$. This model allows for time-varying effects for some covariates and is more flexible than the Fine and Gray model.

The estimation of the time-independent and time-varying coefficients can be obtained based on direct binomial regression (Scheike et al., 2008). Let $\boldsymbol{\eta}(t) = (\boldsymbol{\alpha}(t)^T, \boldsymbol{\gamma}^T)^T$, $D_{\boldsymbol{\eta}} = \partial F_1(t;\mathbf{X})/\partial\boldsymbol{\eta}(t)$ and $w(t,\mathbf{X})$ is some possibly random weight function. Scheike et al. (2008) proposed the following estimating equation:

$$U^*(\boldsymbol{\eta})(t) = \sum_{i=1}^{n} D_{\boldsymbol{\eta}}^T(t;\mathbf{X}_i) w(t,\mathbf{X}_i)\left\{\frac{r_i(t)N_{1i}(t)}{\hat{G}(T_i|\mathbf{X}_i)} - F_1(t;\mathbf{X}_i)\right\}, \tag{4.4.1}$$

where $\hat{G}$, $r_i(t)$ and N_{1i} are the Kaplan-Meier estimator of the survival function for the censoring, the vital status indicator, and the cause 1 counting process as defined for the Fine and Gray model.

4.4.2 Extensions Using Inverse Weighting and the Pattern-Mixture Approach

When the survival data come from a two-stage randomized trial, the subjects who are consistent with the regime of interest should again be weighted to account for those who could have been consistent with the treatment regime but received different treatment due to the second stage randomization. Similar to the extension for the Fine and Gray model, we replace N_{1i} with N_{1i}^w in the estimating equation (4.4.1) which is equivalent to creating an augmented data set as discussed before. The implementation of this weighted method is immediate by using the R function "comp.risk" in the package **timereg** for the Scheike et al. (2008) model on the augmented data. However, obtaining an estimated variance for the CIF is not that

straightforward in this case. Due to the complexity of the model based variance, it is not possible to correct the estimated variance obtained by using the augmented data set. It is also not feasible to use bootstrapping to get an estimated variance for the CIF due to Scheike et al. (2008) model's convergence issues. For these reasons, the simulation results for the inverse weighted Scheike et al. (2008) model will be provided without coverage rates.

We can also think about this model for a two-stage randomized trial as we did in (4.2.6). Then, the CIF estimates for a specific regime can be obtained by splitting the data consistent with the regime into two parts depending on the treatment combinations they received and then fitting the Scheike et al. model to each subset of data using the function "comp.risk" in R package **timereg**. Combining the estimates from these two models as in (4.2.7) will give us the pattern-mixture CIF estimate for the regime of interest. Variance estimates can also be combined as before to obtain a final variance estimate.

4.5 SIMULATION STUDIES

4.5.1 Simulations under the Fine and Gray Model

We conduct simulation studies to compare the performance of our proposed methods and the naïve models in evaluating the covariate effects on the CIFs for two specific regimes $A_1B_1B_1'$ and $A_1B_1B_2'$. Since it is not possible to directly simulate the data such that the cause-1 CIFs for both regimes satisfy the proportional hazards model for sub-distribution in (4.3.1) simultaneously, for the covariates $\mathbf{X} = (X_1, X_2)$ where $X_1 \sim Normal(0, 1)$ and $X_2 \sim Bernoulli(0.5)$, we assume the following model for the type 1 CIF:

$$F_1(t; \mathbf{X}; R; Z_1; Z_2) = 1 - [1 - p\{1 - \exp(-t)\}]^{\exp(\gamma_1 X_1 + \gamma_2 X_2)}, \tag{4.5.1}$$

where $\gamma_1 = R\{Z_1\beta_{11} + (1 - Z_1)\beta_{13}\} + (1 - R)\{Z_2\beta_{12} + (1 - Z_2)\beta_{14}\}$ and $\gamma_2 = R\{Z_1\beta_{21} + (1 - Z_1)\beta_{23}\} + (1 - R)\{Z_2\beta_{22} + (1 - Z_2)\beta_{24}\}$. In other words when patients are treated with A_1B_1, A_1B_1', A_1B_2 and A_1B_2', their CIFs satisfy the Fine and Gray model, and are associated with (X_1, X_2) by (β_{11}, β_{21}), (β_{12}, β_{22}), (β_{13}, β_{23}) and (β_{14}, β_{24}) respectively.

For all sets of simulations, data samples of $n = 400$ were generated for 2000 times. For the two regimes, we followed the probabilities of response and subsequent assignments as specified in scenarios 2 and 3 in Section 2.5. In each setting and each run , first, the two covariates were generated. Then, the response indicator (R), the second assignment indicator for responders (Z_1) and the second assignment indicator for non-responders (Z_2) were generated. Now that it is known which subject receives which treatment after they receive A_1, i.e. $B_1, B_2, B_1^{'}$ or $B_2^{'}$, the inversion of the CIFs can be carried out accordingly to obtain event times and event types as described below. For all settings, the regression coefficients were assumed to be $(\beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}) = (0.3, 0.7, 0.5, 1.0)$ and $(\beta_{21}, \beta_{22}, \beta_{23}, \beta_{24}) = (0.2, 0.6, 0.4, 0.9)$, and $p = 0.4$.

Now we describe how to generate event times T and event types ϵ in detail. Based on n and the simulated R, Z_1 and Z_2, we calculated the numbers of subjects in all treatment options, $n_{A_1B_1}, n_{A_1B_2}, n_{A_1B_1^{'}}$ and $n_{A_1B_2^{'}}$. For A_1B_1, we generated two vectors of uniform random variables U and V of length $n_{A_1B_1}$. Similar vectors of uniform random variables were generated for other treatment options. The inverse of the CIF given in (4.5.1) for A_1B_1 is

$$F^{-1}_{1,A_1B_1}(u; \mathbf{X}) = -\log\left[1 - \frac{1-(1-u)^{\exp(-\beta_{11}X_1 - \beta_{21}X_2)}}{p}\right].$$

For subjects who were treated with A_1B_1, $tmp = 1-(1-U)^{\exp(-\beta_{11}X_1-\beta_{21}X_2)}$ was calculated and when $tmp < p$ the event time was calculated as $T = F^{-1}_{1,A_1B_1}(U; \mathbf{X})$ and the event type was set equal to 1. When $tmp \geq p$, the above inverse function did not exist, implying that the type 2 event occurs before the type 1 event. The conditional distribution of T, given covariates $\mathbf{X}$ and type-2 event occurring first, was modeled as:

$$P(T \leq t | \epsilon = 2, \mathbf{X}) = 1 - \exp\{-t\exp(\beta_{11}X_1 + \beta_{21}X_2)\},$$

so when $tmp \geq p$ the event time was calculated as $T = -\log\left\{(1-V)^{\exp(-\beta_{11}X_1-\beta_{21}X_2)}\right\}$ and event type was set equal to 2. This algorithm has been used in Cheng (2009) and Cheng and Fine (2012), among others. The event times and types for the remaining treatment options were also simulated with their respective uniform random variables and regression coefficients, following the same simulation strategy. Independent censoring was also introduced

to the data set. The censoring time was generated from a $Uniform(2, 5)$ distribution which resulted in $\sim 7\%$ censoring when $P(R = 1) = 0.4$ and $\sim 6\%$ censoring when $P(R = 1) = 0.7$. The observed event time V was then set equal to $\min(T, C)$ and the final event type indicator Δ was set equal to $I(T < C)\epsilon$.

Since we assumed the proportional sub-distribution hazards model for the CIFs of A_1B_1, A_1B_2, $A_1B_1^{'}$, and $A_1B_2^{'}$, the proportional sub-distribution hazards assumption would not hold if we look at the CIFs for the treatment regimes $A_1B_1B_1^{'}$ and $A_1B_1B_2^{'}$. However, the prior work by Shi et al. (2013) suggested that the Fine and Gray model can still provide reliable prediction of the CIF even when the model is misspecified. Hence, in the following tables we list the average values of the estimated CIFs using the methods described in Sections 4.2, 4.3 and 4.4, evaluated at times 0.5 and 1 for $(X_1, X_2) = (1.5, 1)$. Let MS, WMS and PMMS represent the naïve multi-state, inverse weighted multi-state and pattern-mixture multi-state models; FG, WFG and PMFG represent the original Fine and Gray, inverse weighted Fine and Gray and pattern-mixture Fine and Gray models; and SC, WSC and PMSC represent the Scheike et al., inverse weighted Scheike et al. and pattern-mixture Scheike et al. models.

Table 6 contains the results derived from the data sets generated under scenario 2 with $n = 400$. For the first regime, it can be seen that the naïve, inverse weighted and pattern-mixture approaches all perform similarly as expected for both event times and both response rates. However, for the second regime the performance of the naïve approach drops drastically for all three models. The inverse-weighted and the pattern-mixture extension models still perform as desired and provide unbiased estimates of the CIF under all cases. Overall, it can be seen that the WFG and PMFG models perform slightly better than the respective multi-state and Scheike et al. models under this simulation setting. Also, when the inverse-weighted and the pattern-mixture approaches are compared, it can be seen that the pattern-mixture extensions provide slightly better estimates.

Table 7 contains the results derived from the data sets generated under scenario 3 with $n = 400$. For both regimes, the naïve models produce biased estimates of the CIFs at both response rates and both event times. This results in low coverage rates. On the other hand, inverse weighting and pattern-mixture extensions of all models provide much better results

with desirable coverage rates. Again, it can be seen that the WFG and PMFG models perform slightly better than the respective multi-state and Scheike et al. models, when the underlying models for treatment sequences satisfy the Fine and Gray model. Also, generally speaking, it can be seen that the pattern-mixture extensions provide slightly better estimates.

4.5.2 Simulations under the Scheike et al. Model

We also evaluated the performance of the proposed estimators in Sections 4.2-4.4 under the setting where the CIFs of treatment combinations followed Scheike et al. models. To simulate the data, we assumed two covariates $\mathbf{X} = (X_1, X_2)$ had significant effects on the cause 1 CIF with X_1's effect varying with time. The CIFs of the treatment combinations that made up the regimes $A_1B_1B_1'$ and $A_1B_1B_2'$, namely, A_1B_1, A_1B_1', A_1B_2 and A_1B_2', were assumed to follow the models below:

$$
\begin{aligned}
F_{1,A_1B_1}(t;\mathbf{X}) &= 1-[1-p\{1-\exp(-t)\}]^{\exp[I(t\le t_0)\beta_{11}X_1+I(t>t_0)\beta_{21}X_1+\beta_{31}X_2]},\\
F_{1,A_1B_1'}(t;\mathbf{X}) &= 1-[1-p\{1-\exp(-t)\}]^{\exp[I(t\le t_0)\beta_{12}X_1+I(t>t_0)\beta_{22}X_1+\beta_{32}X_2]},\\
F_{1,A_1B_2}(t;\mathbf{X}) &= 1-[1-p\{1-\exp(-t)\}]^{\exp[I(t\le t_0)\beta_{13}X_1+I(t>t_0)\beta_{23}X_1+\beta_{33}X_2]},\\
F_{1,A_1B_2'}(t;\mathbf{X}) &= 1-[1-p\{1-\exp(-t)\}]^{\exp[I(t\le t_0)\beta_{14}X_1+I(t>t_0)\beta_{24}X_1+\beta_{34}X_2]}.
\end{aligned}
$$

To simulate data following these models, first the covariates and response status and randomization indicators were generated. Then, for example for a subject who received treatment combination A_1B_1, two $Uniform$ variables U and V were generated for the two causes. The inverse of the above model for A_1B_1 can be written as:

$$
F^{-1}_{1,A_1B_1}(u;\mathbf{X}) = \begin{cases} -\log\left[1-\frac{1-(1-u)^{\exp(\beta_{11}X_1+\beta_{31}X_2)}}{p}\right], & \text{if } u \le tmp_1,\\ -\log\left[1-\frac{1-(1-u)^{\exp(\beta_{21}X_1+\beta_{31}X_2)}}{p}\right], & \text{if } u > tmp_1, \end{cases}
$$

where $tmp_1 = 1-[1-p\{1-\exp(-t_0)\}]^{\exp(\beta_{11}X_1+\beta_{31}X_2)}$. For each subject the two variables, $tmp_2 = 1-(1-U)^{\exp(\beta_{11}X_1+\beta_{31}X_2)}$ and $tmp_3 = 1-(1-U)^{\exp(\beta_{21}X_1+\beta_{31}X_2)}$ were calculated. Then if $U \le tmp_1$ and $p > tmp_2$, the event time was calculated as $T = F^{-1}_{1,A_1B_1}(U;\mathbf{X})$ and

Table 6: Simulations under Fine and Gray- Scenario-2 with n=400: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation ($\tilde{\sigma}$), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

			$A_1B_1B_1^{'}$					$A_1B_1B_2^{'}$				
t	pr	Model	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov
0.5	0.4	MS	0.47	0.47	0.09	0.09	0.91	0.62	0.70	0.07	0.06	0.68
		WMS		0.47	0.09	0.09	0.91		0.61	0.08	0.06	0.84
		PMMS		0.48	0.09	0.08	0.89		0.61	0.06	0.06	0.90
		FG		0.45	0.09	0.09	0.92		0.70	0.08	0.07	0.77
		WFG		0.45	0.09	0.09	0.92		0.59	0.09	0.09	0.91
		PMFG		0.46	0.09	0.09	0.93		0.62	0.06	0.06	0.94
		SC		0.45	0.10	0.10	0.92		0.71	0.08	0.07	0.72
		WSC		0.45	0.10	–	–		0.61	0.09	–	–
		PMSC		0.47	0.10	0.09	0.91		0.62	0.07	0.06	0.92
	0.7	MS	0.37	0.39	0.09	0.08	0.92	0.45	0.55	0.08	0.07	0.70
		WMS		0.39	0.09	0.08	0.92		0.45	0.09	0.07	0.88
		PMMS		0.39	0.09	0.08	0.91		0.46	0.07	0.07	0.92
		FG		0.36	0.09	0.09	0.92		0.53	0.09	0.09	0.85
		WFG		0.36	0.09	0.09	0.93		0.42	0.09	0.09	0.90
		PMFG		0.37	0.08	0.08	0.93		0.45	0.07	0.07	0.93
		SC		0.36	0.09	0.09	0.92		0.55	0.09	0.09	0.81
		WSC		0.36	0.09	–	–		0.43	0.09	–	–
		PMSC		0.38	0.09	0.08	0.91		0.45	0.08	0.07	0.93
1	0.4	MS	0.64	0.63	0.09	0.08	0.90	0.75	0.82	0.05	0.04	0.52
		WMS		0.63	0.09	0.08	0.90		0.75	0.07	0.05	0.82
		PMMS		0.63	0.08	0.07	0.90		0.72	0.06	0.05	0.83
		FG		0.63	0.10	0.10	0.92		0.86	0.06	0.05	0.45
		WFG		0.63	0.10	0.10	0.92		0.77	0.08	0.08	0.87
		PMFG		0.63	0.09	0.09	0.94		0.75	0.06	0.06	0.91
		SC		0.64	0.11	0.10	0.92		0.87	0.06	0.05	0.37
		WSC		0.64	0.10	–	–		0.78	0.08	–	–
		PMSC		0.64	0.10	0.09	0.92		0.75	0.07	0.06	0.89
	0.7	MS	0.53	0.55	0.09	0.08	0.89	0.59	0.70	0.07	0.06	0.53
		WMS		0.55	0.09	0.08	0.89		0.60	0.08	0.07	0.85
		PMMS		0.54	0.09	0.08	0.90		0.58	0.08	0.07	0.89
		FG		0.53	0.10	0.10	0.93		0.70	0.08	0.08	0.67
		WFG		0.53	0.10	0.10	0.93		0.59	0.09	0.09	0.93
		PMFG		0.53	0.09	0.09	0.93		0.58	0.08	0.08	0.92
		SC		0.53	0.11	0.10	0.92		0.72	0.09	0.08	0.62
		WSC		0.53	0.11	–	–		0.60	0.10	–	–
		PMSC		0.54	0.10	0.09	0.92		0.59	0.09	0.08	0.91

Table 7: Simulations under Fine and Gray- Scenario-3 with n=400: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation ($\tilde{\sigma}$), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

			$A_1B_1B_1^{'}$					$A_1B_1B_2^{'}$				
t	pr	Model	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov
0.5	0.4	MS	0.47	0.42	0.06	0.05	0.85	0.62	0.51	0.06	0.06	0.51
		WMS		0.47	0.06	0.06	0.91		0.60	0.06	0.06	0.90
		PMMS		0.48	0.06	0.06	0.91		0.61	0.05	0.05	0.92
		FG		0.40	0.06	0.06	0.75		0.49	0.07	0.07	0.48
		WFG		0.45	0.06	0.06	0.92		0.59	0.07	0.07	0.92
		PMFG		0.47	0.06	0.06	0.92		0.62	0.05	0.05	0.93
		SC		0.40	0.06	0.06	0.77		0.50	0.07	0.07	0.57
		WSC		0.45	0.07	–	–		0.60	0.07	–	–
		PMSC		0.46	0.07	0.07	0.93		0.62	0.06	0.05	0.92
	0.7	MS	0.37	0.35	0.05	0.05	0.90	0.45	0.39	0.05	0.05	0.70
		WMS		0.38	0.05	0.05	0.92		0.45	0.06	0.05	0.91
		PMMS		0.39	0.05	0.05	0.90		0.46	0.05	0.04	0.92
		FG		0.33	0.05	0.05	0.79		0.36	0.05	0.05	0.55
		WFG		0.36	0.05	0.05	0.92		0.42	0.06	0.06	0.90
		PMFG		0.37	0.06	0.05	0.92		0.45	0.05	0.05	0.94
		SC		0.32	0.05	0.05	0.79		0.36	0.06	0.05	0.60
		WSC		0.36	0.06	–	–		0.43	0.07	–	–
		PMSC		0.37	0.06	0.06	0.91		0.45	0.05	0.05	0.93
1	0.4	MS	0.64	0.59	0.06	0.05	0.80	0.75	0.67	0.06	0.05	0.62
		WMS		0.64	0.06	0.05	0.90		0.74	0.05	0.04	0.90
		PMMS		0.63	0.05	0.05	0.92		0.72	0.04	0.04	0.84
		FG		0.58	0.07	0.06	0.82		0.67	0.07	0.07	0.77
		WFG		0.64	0.07	0.06	0.93		0.76	0.06	0.06	0.92
		PMFG		0.64	0.06	0.06	0.93		0.74	0.04	0.04	0.94
		SC		0.58	0.07	0.07	0.83		0.68	0.07	0.07	0.83
		WSC		0.64	0.07	–	–		0.78	0.06	–	–
		PMSC		0.64	0.06	0.06	0.93		0.74	0.04	0.04	0.94
	0.7	MS	0.53	0.51	0.05	0.05	0.88	0.59	0.54	0.06	0.05	0.81
		WMS		0.54	0.06	0.05	0.90		0.60	0.06	0.05	0.89
		PMMS		0.54	0.05	0.05	0.90		0.58	0.05	0.04	0.92
		FG		0.49	0.06	0.06	0.84		0.52	0.06	0.06	0.79
		WFG		0.53	0.06	0.06	0.93		0.59	0.06	0.06	0.93
		PMFG		0.53	0.06	0.05	0.94		0.58	0.05	0.05	0.93
		SC		0.49	0.06	0.06	0.84		0.53	0.06	0.06	0.80
		WSC		0.53	0.07	–	–		0.60	0.07	–	–
		PMSC		0.53	0.06	0.06	0.93		0.58	0.05	0.05	0.94

event type was set equal to 1. Similarly, if $U > tmp_1$ and $p > tmp_3$, the event time was calculated as $T = F^{-1}_{1,A_1B_1}(U;\mathbf{X})$ and event type was set equal to 1.

For the case where $U \leq tmp_1$ but $p \leq tmp_2$, the CIF was modeled as:

$$P(T \leq t|\epsilon = 2, \mathbf{X}) = 1 - \exp\{-t\exp(\beta_{11}X_1 + \beta_{31}X_2)\},$$

and to obtain the event time it was inverted as $T = -\log\left\{(1-V)^{\exp(-\beta_{11}X_1 - \beta_{31}X_2)}\right\}$ and the event type was set equal to 2. Similarly, for the case where $U > tmp_1$ but $p \leq tmp_3$ the CIF was modeled as:

$$P(T \leq t|\epsilon = 2, \mathbf{X}) = 1 - \exp\{-t\exp(\beta_{21}X_1 + \beta_{31}X_2)\},$$

and to obtain the event time it was inverted as $T = -\log\left\{(1-V)^{\exp(-\beta_{21}X_1 - \beta_{31}X_2)}\right\}$ and the event type was again set equal to 2. For the other three treatment combinations similar data generation procedures were followed. The regression coefficients associated with X_1 were chosen as $(\beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}) = (0.3, 0.7, 0.5, 1.0)$, $(\beta_{21}, \beta_{22}, \beta_{23}, \beta_{24}) = (0.5, 0.9, 0.7, 1.2)$ and the ones associated with X_2 were set as $(\beta_{31}, \beta_{32}, \beta_{33}, \beta_{34}) = (0.2, 0.6, 0.4, 0.9)$ where $p = 0.4$. The time point where the effect of X_1 was assumed to change, t_0, was set equal to 0.75. Independent censoring was introduced to the data set by generating the censoring time C from a $Uniform(2, 4)$ distribution and letting $V = \min(T, C)$ and $\Delta = I(T < C)\epsilon$ resulted in $\sim$ 10% censoring for $P(R = 1) = 0.4$ and $\sim$ 9% censoring for $P(R = 1) = 0.7$. The data generation process was repeated for cases $P(R = 1) = 0.4$ and $P(R = 1) = 0.7$ and for scenarios 2 and 3 in Section 2.5. In the following tables we list the average estimated CIF values evaluated at times 0.5 and 1 for $(X_1, X_2) = (1.5, 1)$.

Again, we let MS, WMS and PMMS represent the naïve multi-state, inverse weighted multi-state and pattern-mixture multi-state models; FG, WFG and PMFG represent the naïve Fine and Gray, inverse weighted Fine and Gray and pattern-mixture Fine and Gray models; and SC, WSC and PMSC represent the Scheike et al., inverse weighted Scheike et al. and pattern-mixture Scheike et al. models.

Table 8 is a summary of the results obtained under scenario 2. For the first regime it can be seen that the pattern-mixture and inverse weighted methods perform similar to the naïve

method for both event times and both response rates as expected. However for the second regime, the naïve models fail to generate unbiased estimates for the CIF and fail to achieve desirable coverage rates. The proposed extensions on FG and SC models all perform much better for both event times and both response rates. Overall it can be seen that extensions of SC models produce estimates that are slightly closer to the true values, when the underlying CIFs for treatment combinations follow the Scheike et al. model.

Table 9 is a summary of the results obtained under scenario 3. Under this scenario the comparison of the methods for the two specific time points are different. For $t = 0.5$ it can be seen that the naïve MS model is performing slightly better compared to the extensions, the FG model performs similar to the extensions, and WSC and PMSC models perform much better compared to the naïve SC model. However, as we look at the second time point $t = 1$ and as a result we include more data points in the analysis, it can be seen that for all models the extensions work much better for both regimes under all conditions compared to the original models.

Looking at the results from both scenarios, we can say that when the data is suspected to involve a covariate with a time-varying effect like under the Scheike et al. model, it is much safer to fit the WSC and PMSC models with a slight preference towards PMSC if there are adequate data in each stratum. Even if one is not aware of the time-varying effect, the proposed extensions of MS and FG models are still much safer alternatives compared to the naïve approaches.

4.6 SUMMARY

In this chapter we proposed two extensions, inverse weighting and pattern-mixture, on three existing regression models on the CIF which can be used when dynamic treatment regimes are of interest and a two-stage randomization is carried out. The inferences for the proposed extension models are simple and the analysis can be carried out by slightly manipulating existing R packages. We showed that the proposed extension models produce more reliable

Table 8: Simulations under Scheike et al.- Scenario-2 with n=400: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation ($\tilde{\sigma}$), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

			$A_1B_1B_1^{'}$					$A_1B_1B_2^{'}$				
t	pr	Model	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov
0.5	0.4	MS	0.47	0.52	0.09	0.09	0.89	0.62	0.73	0.06	0.06	0.51
		WMS		0.52	0.09	0.09	0.89		0.64	0.08	0.06	0.83
		PMMS		0.52	0.09	0.08	0.85		0.64	0.06	0.06	0.91
		FG		0.50	0.10	0.09	0.91		0.74	0.07	0.07	0.58
		WFG		0.50	0.10	0.09	0.91		0.64	0.09	0.08	0.90
		PMFG		0.52	0.09	0.09	0.89		0.65	0.06	0.06	0.91
		SC		0.47	0.12	0.11	0.91		0.72	0.09	0.08	0.71
		WSC		0.47	0.12	–	–		0.62	0.10	–	–
		PMSC		0.49	0.12	0.11	0.88		0.63	0.09	0.07	0.89
	0.7	MS	0.37	0.44	0.09	0.08	0.88	0.45	0.59	0.08	0.07	0.52
		WMS		0.44	0.09	0.08	0.88		0.50	0.08	0.07	0.86
		PMMS		0.44	0.08	0.08	0.85		0.50	0.07	0.07	0.90
		FG		0.42	0.09	0.09	0.93		0.58	0.08	0.08	0.67
		WFG		0.42	0.09	0.09	0.93		0.48	0.09	0.09	0.93
		PMFG		0.43	0.08	0.08	0.92		0.49	0.07	0.07	0.91
		SC		0.38	0.11	0.10	0.91		0.56	0.10	0.10	0.79
		WSC		0.38	0.11	–	–		0.45	0.10	–	–
		PMSC		0.39	0.11	0.11	0.90		0.46	0.09	0.09	0.91
1	0.4	MS	0.73	0.70	0.08	0.07	0.90	0.80	0.85	0.04	0.04	0.68
		WMS		0.70	0.08	0.07	0.90		0.79	0.06	0.04	0.85
		PMMS		0.69	0.08	0.07	0.86		0.75	0.06	0.05	0.80
		FG		0.72	0.09	0.09	0.92		0.90	0.05	0.04	0.42
		WFG		0.72	0.09	0.09	0.91		0.82	0.07	0.07	0.85
		PMFG		0.71	0.08	0.08	0.93		0.79	0.06	0.06	0.89
		SC		0.74	0.11	0.10	0.90		0.93	0.05	0.04	0.29
		WSC		0.74	0.11	–	–		0.85	0.08	–	–
		PMSC		0.74	0.10	0.09	0.90		0.80	0.08	0.07	0.85
	0.7	MS	0.63	0.62	0.08	0.08	0.92	0.67	0.75	0.06	0.05	0.63
		WMS		0.62	0.08	0.08	0.92		0.67	0.08	0.06	0.86
		PMMS		0.61	0.08	0.08	0.90		0.65	0.08	0.07	0.90
		FG		0.62	0.10	0.09	0.92		0.77	0.07	0.07	0.64
		WFG		0.62	0.10	0.09	0.92		0.68	0.09	0.09	0.90
		PMFG		0.61	0.09	0.09	0.92		0.65	0.08	0.08	0.91
		SC		0.64	0.11	0.11	0.91		0.80	0.09	0.08	0.57
		WSC		0.64	0.11	–	–		0.69	0.10	–	–
		PMSC		0.64	0.11	0.10	0.90		0.67	0.10	0.09	0.90

Table 9: Simulations under Scheike et al.- Scenario-3 with n=400: specific time point (t), probability of response (pr), true cumulative incidence (True), mean of estimates (Mean), empirical standard deviation ($\tilde{\sigma}$), mean of estimated standard deviations ($\hat{\sigma}$), coverage rate of 95% confidence intervals (Cov)

			$A_1B_1B_1^{'}$					$A_1B_1B_2^{'}$				
t	pr	Model	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov	True	Mean	$\tilde{\sigma}$	$\hat{\sigma}$	Cov
0.5	0.4	MS	0.47	0.47	0.06	0.06	0.95	0.62	0.56	0.06	0.06	0.76
		WMS		0.51	0.06	0.06	0.85		0.64	0.06	0.05	0.89
		PMMS		0.52	0.06	0.06	0.82		0.64	0.05	0.04	0.88
		FG		0.45	0.06	0.06	0.94		0.54	0.07	0.06	0.76
		WFG		0.50	0.06	0.06	0.91		0.64	0.07	0.06	0.93
		PMFG		0.52	0.06	0.06	0.85		0.65	0.05	0.05	0.84
		SC		0.41	0.07	0.07	0.84		0.52	0.08	0.07	0.69
		WSC		0.46	0.08	–	–		0.62	0.08	–	–
		PMSC		0.46	0.08	0.07	0.93		0.62	0.06	0.06	0.90
	0.7	MS	0.37	0.40	0.05	0.05	0.93	0.45	0.44	0.05	0.05	0.92
		WMS		0.43	0.05	0.05	0.77		0.49	0.06	0.05	0.84
		PMMS		0.44	0.05	0.05	0.73		0.50	0.05	0.04	0.75
		FG		0.38	0.05	0.05	0.95		0.42	0.05	0.05	0.87
		WFG		0.42	0.05	0.05	0.88		0.48	0.05	0.06	0.93
		PMFG		0.43	0.05	0.05	0.82		0.49	0.05	0.05	0.81
		SC		0.33	0.06	0.06	0.86		0.37	0.06	0.06	0.72
		WSC		0.37	0.07	–	–		0.44	0.07	–	–
		PMSC		0.38	0.07	0.06	0.92		0.45	0.06	0.05	0.94
1	0.4	MS	0.73	0.65	0.05	0.05	0.64	0.80	0.72	0.05	0.04	0.55
		WMS		0.70	0.05	0.05	0.90		0.79	0.04	0.04	0.89
		PMMS		0.69	0.05	0.04	0.86		0.76	0.04	0.03	0.74
		FG		0.66	0.06	0.06	0.79		0.74	0.06	0.06	0.84
		WFG		0.72	0.06	0.06	0.95		0.82	0.05	0.05	0.89
		PMFG		0.72	0.05	0.05	0.95		0.79	0.04	0.04	0.92
		SC		0.67	0.07	0.07	0.87		0.76	0.07	0.07	0.90
		WSC		0.73	0.07	–	–		0.85	0.06	–	–
		PMSC		0.73	0.06	0.06	0.94		0.80	0.04	0.04	0.93
	0.7	MS	0.63	0.58	0.05	0.05	0.82	0.67	0.61	0.05	0.05	0.78
		WMS		0.62	0.05	0.05	0.92		0.67	0.05	0.04	0.90
		PMMS		0.61	0.05	0.05	0.92		0.65	0.04	0.04	0.90
		FG		0.58	0.06	0.06	0.84		0.61	0.06	0.06	0.84
		WFG		0.62	0.06	0.06	0.94		0.67	0.06	0.06	0.93
		PMFG		0.62	0.05	0.05	0.94		0.65	0.05	0.05	0.93
		SC		0.58	0.07	0.06	0.88		0.62	0.07	0.06	0.87
		WSC		0.63	0.07	–	–		0.69	0.07	–	–
		PMSC		0.63	0.06	0.06	0.94		0.66	0.05	0.05	0.92

estimates for the CIF under a two-stage randomization setting for various conditions using simulation. In general the pattern-mixture extensions of all three models perform slightly better so they should be preferred over the inverse weighted extensions when there are adequate data in each subgroup. We conclude that when there are time-varying effects on the CIF it is safer to fit the extensions of the Scheike et al. model. Overall, regardless of the nature of data, the proposed extensions provide a more reliable way to analyze covariate effects on the CIF of a dynamic treatment regime and should be preferred over the existing naïve models.

5.0 REMARKS AND FUTURE WORK

In this dissertation, we focused on two-stage dynamic treatment regimes under the presence of competing risks. The methods provide unbiased estimation of the CIF which facilitates the comparison of dynamic treatment regimes and, as a result, helps choose the best regime that results in minimal chance of experiencing the failure of interest. Both inverse probability weighting and pattern-mixture estimators work well, and we provide explicit variance estimators which make inference straightforward. There is not much difference between the two approaches except when there are inadequate number of subjects receiving a specific treatment combination which may make the pattern-mixture estimator unreliable if it is calculable. Other than that, the choice simply depends on researchers' familiarity with the two approaches.

We also provide methods to extend existing regression models on the CIF to dynamic treatment regimes. They will be useful in directly evaluating covariate effects on the CIF for a specific regime. The extended models can easily be applied to any data set using existing packages. Both extensions work much better than naïvely fitting a model on all the subjects who are consistent with a specific regime with pattern-mixture extensions working slightly better. This is expected. The models considered in Chapter 4 are more likely to hold on the treatment sequence level than on the regime level which is more heterogeneous regarding the mixture of responders and non-responders.

In this dissertation, we have focused on survival outcomes subject to competing-risk censoring. The methods provided in this dissertation will also be useful even when there are no competing risks. The same approaches can be taken on extending regression models for typical survival outcomes and more general continuous outcomes. Developing such models will

be very useful for comparison of dynamic treatment regimes and evaluating covariate effects on a specific regime with respect to different types of outcomes. It may also be desirable to develop model selection procedures that would work under a two-stage randomization to help choose the most effective covariates on these outcomes. These may be future research topics.

BIBLIOGRAPHY

Aalen, O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics 6*, 701–726.

Andersen, P. K., S. Rosthoj, S. Rosthoej, S. Z. Abildstrom, and S. Rosthøj (2002). Competing risks as a multi-state model. *Statistical Methods in Medical Research 11*(2), 203–215.

Breslow, N. (1974). Covariance analysis of censored survival data. *Biometrics 30*, 89–99.

Cheng, S. C., J. P. Fine, and L. J. Wei (1998). Prediction of cumulative incidence function under the proportional hazards model. *Biometrics 54*, 219–228.

Cheng, Y. (2009). Modeling cumulative incidences of dementia and dementia-free death using a novel three-parameter logistic function. *The International Journal of Biostatistics*, DOI: 10.2202/1557-4679.1183.

Cheng, Y. and J. P. Fine (2012). Cumulative incidence association models for bivariate competing risks data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 74*, 183–202.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B: Methodological 34*, 187–220.

Dawson, R. and P. W. Lavori (2010). Sample size calculations for evaluating treatment policies in multi-stage designs. *Clinical Trials 7*, 643–652.

Fine, J. P. (2001). Regression modeling of competing crude failure probabilities. *Biostatistics (Oxford) 2*(1), 85–97.

Fine, J. P. and R. J. Gray (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association 94*, 496–509.

Fleming, T. R. and D. P. Harrington (1991). *Counting Processes and Survival Analysis*. John Wiley & Sons.

Goldberg, Y. and M. R. Kosorok (2012). Q-learning with censored data. *Annals of Statistics 40*(1), 529–560.

Gooley, T. A., W. Leisenring, J. Crowley, and B. E. Storer (1999). Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Statistics in Medicine 18*, 695–706.

Gray, R. J. (1988). A class of K-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics 16*, 1140–1154.

Guo, X. and A. Tsiatis (2005). A weighted risk set estimator for survival distributions in two-stage randomization designs with censored survival data. *The International Journal of Biostatistics 1*(1), 1–17.

Henderson, R., P. Ansell, and D. Alshibani (2010). Regret-regression for optimal dynamic treatment regimes. *Biometrics 66*(4), 1192–1201.

Hyun, S., Y. Sun, and R. Sundaram (2009). Assessing cumulative incidence functions under the semiparametric additive risk model. *Statistics in Medicine 28*(22), 2748–2768.

Jeong, J. and J. P. Fine (2007). Parametric regression on cumulative incidence function. *Biostatistics 55*, 184–200.

Jin, Z., D. Y. Lin, L. J. Wei, and Z. Ying (2003). Rank-based inference for the accelerated failure time model. *Biometrika 90*(2), 341–353.

Kalbfleisch, J. D. and R. L. Prentice (2002). *The Statistical Analysis of Failure Time Data*. New York: Wiley.

Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association 53*, 457–481.

Klein, J. P. (2006). Modelling competing risks in cancer stuties. *Statistics in Medicine 25*(6), 1015–1034.

Klein, J. P. and P. K. Andersen (2005). Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics 61*(1), 223–229.

Klein, J. P. and M. L. Moeschberger (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer-Verlag Inc.

Koller, M., H. Raatz, E. Steyerberg, and M. Wolbers (2012). Competing risks and the clinical community: irrelevance or ignorance? *Statistics in medicine 31*, 1089–1097.

Korn, E. L. and F. J. Dorey (1992). Applications of crude incidence curves. *Statistics in Medicine 11*, 813–829.

Lavori, P. and R. Dawson (2004). Dynamic treatment regimes: practical design considerations. *Clinical trials 1*, 9–20.

Lavori, P. W. and R. Dawson (2000). A design for testing clinical strategies: Biased adaptive within-subject randomization. *Journal of the Royal Statistical Society, Series A: Statistics in Society 163*(1), 29–38.

Lin, D. Y. (1997). Non-parametric inference for cumulative incidence functions in competing risks studies. *Statistics in Medicine 16*, 901–910.

Lin, D. Y., T. R. Fleming, and L. J. Wei (1994). Confidence bands for survival curves under the proportional hazards model. *Biometrika 81*, 73–81.

Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association 88*, 125–134.

Lokhnygina, Y. and J. D. Helterbrand (2007). Cox Regression Methods for Two-Stage Randomization Designs. *Biometrics 63*(2), 422–428.

Lunceford, J. K., M. Davidian, and A. A. Tsiatis (2002). Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials. *Biometrics 58*(1), 48–57.

Miyahara, S. and A. S. Wahed (2010). Weighted kaplan-meier estimators for two-sage treatment regimes. *Statistics in Medicine 29*(1), 2581–2591.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Series B: Statistical Methodology 65*(2), 331–366.

Murphy, S. A. (2005). An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine 24*(10), 1455–1481.

Murphy, S. A. and D. Bingham (2009). Screening Experiments for Developing Dynamic Treatment Regimes. *Journal of the American Statistical Association 104*(485), 391–408.

Murphy, S. A., K. G. Lynch, D. Oslin, J. R. McKay, and T. T. (2007). Developing adaptive treatment strategies in substance abuse research. *Drug and Alcohol Dependence 88S*, S24–30.

Murphy, S. A., M. J. van der Laan, and J. M. Robins (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association 96*(456), 1410–1423.

Robins, J. (1986). A new approach to causal inference in mortality studies with sustained exposure periods-application to control of the healthy survivor effect. *Mathematical Modelling 7*, 1393–1512.

Robins, J. (1987). Addendum to "a new approach to causal inference in mortality studies with sustained exposure periods-application to control of the healthy survivor effect". *Computers and Mathematics with Applications 14*, 923–945.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association 89*, 846–866.

Scheike, T. H., M.-J. Zhang, and T. A. Gerds (2008). Predicting cumulative incidence probability by direct binomial regression. *Biometrika 95*(1), 205–220.

Shi, H., Y. Cheng, and J.-h. Jeong (2013). Constrained parametric model for simultaneous inference of two cumulative incidence functions. *Biometrical Journal 55*, 82–96.

Stone, R. M., D. T. Berg, S. L. George, R. K. Dodge, P. A. Paciucci, P. Schulman, E. J. Lee, J. O. Moore, B. L. Powell, and C. A. Schiffer (1995). Granulocyte-macrophage colony-stimulating factor after initial chemotherapy for elderly patients with primary acute myelogenous leukemia. *The New England Journal of Medicine 332*(25), 1671–1677.

Thall, P., R. Millikan, and S. H.C. (2000). Evaluating multiple treatment courses in clinical trials. *Statistics in Medicine 19*, 1011–1028.

Wahed, A. S. (2010). Inference for two-stage adaptive treatment strategies using mixture distributions. *Journal of the Royal Statistical Society 59*, 1–18.

Wahed, A. S. and X. Tang (2013). Pattern-mixture-type estimation and testing of neuroblastoma treatment regimes. Statistics in Medicine, in press.

Wahed, A. S. and A. A. Tsiatis (2006). Semiparametric efficient estimation of survival distributions in two-stage randomisation designs in clinical trials with censored data. *Biometrika 93*(1), 163–177.

Wei, L. J., Z. Ying, and D. Y. Lin (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika 77*, 845–851.

Zhang, M.-J. and J. Fine (2008). Summarizing differences in cumulative incidence functions. *Statistics in Medicine 27*(24), 4939–4949.

Zhao, Y., M. R. Kosorok, and D. Zeng (2009). Reinforcement learning design for cancer clinical trials. *Statistics in Medicine 28*(26), 3294–3315.